



# INTERCONNECTS ARCHITECTURES FOR MANY-CORE ERA USING SURFACE-WAVE COMMUNICATION

Ammar J. M. Karkar

A Thesis Submitted for the Degree of  
Doctor of Philosophy at Newcastle University

School of Electrical and Electronic Engineering  
Faculty of Science, Agriculture and Engineering

August 2016



## DECLARATION

---

I hereby declare that this thesis is my own work and effort and that it has not been submitted anywhere for any award. Where other sources of information have been used, they have been acknowledged.

*Newcastle upon Tyne, August 2016*

---

Ammar Karkar

## CERTIFICATE OF APPROVAL

---

I confirm that, to the best of my knowledge, this thesis is from the student's own work and effort, and all other sources of information used have been acknowledged. This thesis has been submitted with my approval.

---

ALEX YAKOVLEV



*To my greatest supporters that is my wonderful parents, my beloved wife,  
Ghsaq, and my lovely daughter, Maryam.*  
— Ammar

## ACKNOWLEDGEMENTS

---

I would like to express my sincere gratitude to my supervisors Prof. Alex Yakovlev and Prof. Terrence Mak for their support and guidance through my studies. They have been and always will be a source of inspiration and my role model as a researcher.

I am also grateful to my colleagues and friends in the School of Electrical and Electronic Engineering, especially those in Microelectronic Systems research group, at Newcastle University for their assistance and guidance in my studies. Especially, I appreciate the support of my wonderful friends and colleagues Dr. Nizar Dahir, Dr. Ra'ed Aldujaily, Dr. Ahmed Sabaawi, Dr. Walled Amer and Dr. Hussein Leftah for fruitful discussions, productive suggestions, and subjective criticism. In addition, I would like to thank Dr. Graeme Coapes for his help and advice regarding the building of the SNN benchmark.

I would also wish to thank Prof. Kin-Fai Tong and Dr. Janice Turner for their help in offering their experiment results and knowledge, which assist me toward the understanding of Zennek surface wave.

I am also thankful for Dr. Xiaohang Wang for his assistance in building application benchmarks.

I am also grateful to Dr Maurizio Palesi for his advice in modifying the NoC simulator to include the feature of simulating the VCs.

Finally, I would like to offer my special regards to all the staff of the school of Electrical and Electronic Engineering in Newcastle university.

## ABSTRACT

---

Networks-on-chip (NoCs) is a communication paradigm that has emerged aiming to address on-chip communication challenges and to satisfy interconnection demands for chip-multiprocessors (CMPs). Nonetheless, there is continuous demand for even higher computational power, which is leading to a relentless downscaling of CMOS technology to enable the integration of many-cores. However, technology downscaling is in favour of the gate nodes over wires in terms of latency and power consumption. Consequently, this has led to the era of many-core processors where power consumption and performance are governed by inter-core communications rather than core computation. Therefore, NoCs need to evolve from being merely metal-based implementations which threaten to be a performance and power bottleneck for many-core efficiency and scalability.

To overcome such intensified inter-core communication challenges, this thesis proposes a novel interconnect technology: the surface-wave interconnect (SWI). This new RF-based on-chip interconnect has notable characteristics compared to cutting-edge on-chip interconnects in terms of CMOS compatibility, high speed signal propagation, low power dissipation, and massive signal fan-out. Nonetheless, the realization of the SWI requires investigations at different levels of abstraction, such as the device integration and RF engineering levels. The aim of this thesis is to address the networking and system level challenges and highlight the potential of this interconnect. This should encourage further research at other levels of abstraction. Two specific system-level challenges crucial in future many-core systems are tackled in this study, which are cross-the-chip global communication and one-to-many communication.

This thesis makes four major contributions towards this aim. The first is reducing the NoC average-hop count, which would otherwise increase packet-latency exponentially, by proposing a novel hybrid interconnect architecture. This hybrid architecture can not only utilize both regular metal-wire and SWI, but also exploits merits of both bus and NoC architectures in terms of connectivity compared to other general-purpose on-chip interconnect architectures. The second contribution addresses global communication issues by developing a distance-based weighted-round-robin arbitration (DWA) algorithm. This technique prioritizes global communication to be send via SWI short-cuts, which offer more efficient power dissipation and faster across-the-chip signal propagation. Results obtained using a cycle-accurate simulator demonstrate the effectiveness of the proposed system architecture in terms of significant power reduction, consider-

able average delay reduction and higher throughput compared to a regular NoC. The third contribution is in handling multicast communications, which are normally associated with traffic overload, hotspots and deadlocks and therefore increase, by an order of magnitude the power consumption and latency. This has been achieved by proposing a novel routing and centralized arbitration schemes that exploits the SWI's remarkable fan-out features. The evaluation demonstrates drastic improvements in the effectiveness of the proposed architecture in terms of power consumption ( 2-10x) and performance ( 22x) but with negligible hardware overheads ( 2%). The fourth contribution is to further explore multicast contention handling in a flexible decentralized manner, where original techniques such as stretch-multicast and ID-tagging flow control have been developed. A comparison of these techniques shows that the decentralized approach is superior to the centralized approach with low traffic loads, while the latter outperforms the former near and after NoC saturation.

## PUBLICATIONS

---

### *Journal and magazines publications:*

1. **Ammar J. M. Karkar**; Janice E. Turner; Kenneth Tong; Ra'ed Al-Dujaily; Terrence Mak ; Alex Yakovlev; Fei Xia, *Hybrid wire-surface wave interconnects for next-generation networks-on-chip* , IET Computers and Digital Techniques, 2013, 7, (6), p. 294-303, DOI: 10.1049/iet-cdt.2013.0030, IET Digital Library.
2. **Ammar J. M. Karkar**; Terrence Mak; N. Dahir ; Ra'ed Al-Dujaily; Kenneth Tong; Alex Yakovlev; *Network-on-Chip Multicast Architectures Using Hybrid Wire and Surface-Wave Interconnects*, 2016, IEEE Transactions on Emerging topics in Computing, special issue on Emerging Computational paradigms and Architectures for Multi-core Platforms. ISSN 2168-6750. doi: 10.1109/TETC.2016.2551043
3. **Ammar J. M. Karkar**, T. Mak, K. Tong, and A. Yakovlev, *A Survey of Emerging Interconnects for On-chip Efficient Multicast and Broadcast in Many-cores*, Circuits and Systems Magazine, IEEE, 16(1):58-72, Firstquarter 2016. ISSN 1531-636X. doi: 10.1109/M-CAS.2015.2510199.

### *Conference publications:*

1. **Ammar J. M. Karkar**; N. Dahir; R. Al-Dujaily; K. Tong; T. Mak ; A. Yakovlev , *Hybrid wire-surface wave architecture for one-to-many communication in networks-on-chip*, Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014 , pp.1,4, 24-28 March 2014, doi: 10.7873/DATE.2014.287
2. **Ammar J. M. Karkar**, K. Tong, T. Mak, and A. Yakovlev, *Mixed wire and surfacewave communication fabrics for decentralized on-chip multicasting*, in Design, Automation and Test in Europe Conference and Exhibition (DATE), 2015, (EDA Consortium), pp 794-799, March 2015. ISBN 978-3-9815370-4-8.

### *Workshop and forum publications:*

1. **Ammar J. M. Karkar**, R. Al-Dujaily, A. Yakovlev, K. Tong, and T. Mak, *Surface wave communication system for on-chip and off-chip interconnects*, in Proceedings of the Fifth International Workshop

on Network on Chip Architectures, NoCArc'12, (New York, NY, USA), pp. 11-16, ACM, 2012.

2. **Ammar J. M. Karkar**, T. Mak and A. Yakovlev, *Surface wave communication systems for on-chip and off chip interconnects*, Proceedings of UK Electronics Forum (UKEF'12), Newcastle, 30-31th Aug. 2012, pp. 59- 64.

*I also contributed in the following works:*

1. Mengyuan Wu; **Ammar J. M. Karkar**; Bo Liu; A. Yakovlev; G. Gielen; V. Grout, *Network on Chip optimization based on surrogate model assisted evolutionary algorithms*, Evolutionary Computation (CEC), 2014 IEEE Congress on , vol., no., pp.3266,3271, 6-11 July 2014 doi: 10.1109/CEC.2014.6900559.
2. B. Liu, F. Fernandez, G. Gielen, **Ammar J. M. Karkar**, A. Yakovlev, V. Grout, *SMAS: A Generalized and Efficient Framework for Computationally Expensive Electronic Design Optimization Problems* , Computational Intelligence in Analog and Mixed-Signal (AMS) and Radio-Frequency (RF) Circuit Design, Springer, 2015 (Accepted).

## CONTENTS

---

<b>I</b>	<b>Thesis Chapters</b>	<b>1</b>
<b>1</b>	<b>INTRODUCTION</b>	<b>2</b>
1.1	Motivation . . . . .	2
1.2	Statement of originality . . . . .	5
1.3	Thesis Organization . . . . .	7
<b>2</b>	<b>BACKGROUND AND LITERATURE REVIEW</b>	<b>10</b>
2.1	Introduction . . . . .	10
2.2	Background . . . . .	10
2.2.1	Networks-on-chip . . . . .	10
2.2.2	Deadlock, Livelock and Starvation . . . . .	14
2.2.3	Flow Control . . . . .	16
2.2.4	Routing Algorithms . . . . .	19
2.2.5	Arbitration and Allocation . . . . .	21
2.2.6	Multi/Many-Core Processors . . . . .	23
2.2.7	Cache Coherence . . . . .	24
2.2.8	Three-Dimensional Integration . . . . .	25
2.3	Literature Review . . . . .	26
2.3.1	Current and Emerging Interconnects . . . . .	26
2.3.2	Existing NoC Architectures . . . . .	40
2.3.3	NoC Simulators and Models . . . . .	42
<b>3</b>	<b>SURFACE-WAVE ON-CHIP INTERCONNECT</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Surface-Wave interconnect fabric . . . . .	46
3.2.1	Surface Wave Interconnects (SWI) implementation	46
3.2.2	SWI Links design . . . . .	47
3.2.3	SWI Challenges . . . . .	49
3.3	Zenneck Surface wave modelling . . . . .	49
3.3.1	Analysis of link power dissipation . . . . .	49
3.3.2	Communication system performance . . . . .	50
3.4	Experiment Results . . . . .	53
3.5	Summary and Conclusion . . . . .	55
<b>4</b>	<b>HYBRID WIRE AND SURFACE-WAVE ARCHITECTURE FOR ON-CHIP GLOBAL COMMUNICATIONS</b>	<b>57</b>
4.1	Introduction . . . . .	57
4.2	Relevant Work on Cross-the-chip Communications . .	58
4.2.1	Multi-hop Challenges and Wire-based Solutions	58
4.2.2	Routing in Hybrid Architectures . . . . .	58

4.3	Hybrid Wire and Surface-wave Interconnect Architecture	59
4.3.1	Addressing Multi-hop Challenges Using W-SWI Architecture . . . . .	59
4.3.2	Routing scheme . . . . .	62
4.3.3	Distance-based Weighted-round-robin Arbitration (DWA) Algorithm . . . . .	63
4.4	System level evaluation and discussion . . . . .	66
4.4.1	Performance Evaluation . . . . .	68
4.4.2	Power Consumption . . . . .	70
4.4.3	Area Estimation . . . . .	71
4.5	Summary and Conclusion . . . . .	73
5	WIRE AND SURFACE-WAVE ARCHITECTURE WITH CENTRALIZED CONTROL FOR MULTICAST	74
5.1	Introduction and Motivation . . . . .	74
5.1.1	Motivation . . . . .	75
5.2	Related Work in Current and Emerging Interconnects for Multicast . . . . .	77
5.2.1	Wire-based Multicast Routing and Architectures	77
5.2.2	Optical interconnects . . . . .	79
5.2.3	Wireless Interconnects . . . . .	80
5.2.4	Transmission Lines . . . . .	81
5.2.5	SWI Fanout Feature . . . . .	81
5.3	W-SWI multicast routing scheme . . . . .	83
5.4	Proposed Arbitration and Allocation Schemes . . . . .	84
5.4.1	Contention Challenges . . . . .	84
5.4.2	Centralized Arbitration . . . . .	85
5.4.3	Communication Protocol . . . . .	90
5.5	System Level Evaluation and Discussion . . . . .	91
5.5.1	Performance Improvements . . . . .	92
5.5.2	Power Reduction . . . . .	94
5.5.3	Quality Of Service (QoS) . . . . .	95
5.5.4	Comparison with Related Work . . . . .	96
5.5.5	W-SWI-C for Spiking Neural Network (SNN) .	97
5.5.6	Area Overhead Evaluation . . . . .	98
5.6	Summary and Conclusion . . . . .	99
6	WIRE AND SURFACE-WAVE ARCHITECTURE WITH DECENTRALIZED CONTROL FOR MULTICAST	101
6.1	Introduction and Motivation . . . . .	101
6.2	Decentralized Arbitration and allocation . . . . .	102
6.2.1	Stretched Multicast . . . . .	103
6.2.2	Deadlock-free Flow Control . . . . .	104
6.2.3	Communication Protocol . . . . .	106
6.3	System Level Evaluation and Discussion . . . . .	108
6.3.1	Performance Improvements . . . . .	110
6.3.2	Power Reduction . . . . .	113



6.3.3	Evaluation with Real Application Benchmark . . . . .	113
6.3.4	Area Overhead Evaluation . . . . .	115
6.4	Summary and Conclusion . . . . .	117
7	CONCLUSIONS AND FUTURE WORK . . . . .	118
7.1	Summary and Conclusion . . . . .	118
7.2	Future Work . . . . .	120
<b>II</b>	<b>Thesis Appendices . . . . .</b>	<b>121</b>
A	NOXIM SIMULATOR IMPROVEMENTS . . . . .	122
A.1	SWI Channel Modelling . . . . .	122
A.2	Virtual Channel Modelling . . . . .	122
A.3	1-to-M Traffic Modelling . . . . .	123
<b>III</b>	<b>Thesis Bibliography . . . . .</b>	<b>125</b>
	BIBLIOGRAPHY . . . . .	126

## LIST OF FIGURES

Figure 1.1	Processing element number is projected to scale exponentially according to ITRS system-on-chip (SoC) [14]. . . . .	3
Figure 1.2	(a) 1-to-M, M-to-1 and 1-to-1 traffic percentage in different CMP real benchmark when a broadcast-based cache coherence protocol (token coherence) is used [16], (b) NoC simulation shows an increase in average delay and fast network saturation when broadcast treated as unicast in a $6 \times 4$ NoC. . . . .	4
Figure 1.3	Thesis organization. . . . .	8
Figure 2.1	Examples of on-chip communication network topologies. . . . .	12
Figure 2.2	main components of network-on-chip (NoC) such as router(R) and network interface (NI) common micro-architecture. . . . .	14
Figure 2.3	Example scenario of a deadlock resulting from cyclic channel dependency. . . . .	15
Figure 2.4	illustration of segmentation and the encapsulation process of message into packets, flits, and phits depending on the NoC resources and the flow control techniques. . . . .	16
Figure 2.5	Example scenario showing the benefit of using virtual channels (VCs) on mitigating blocking channels issues, even though the buffer size remain the same. . . . .	18
Figure 2.6	illustration of the allowed (solid arrows) and prohibited (dashed arrows) turns in different turn-model routing algorithms in two-dimensional (2D) mesh networks-on-chip (NoCs) . . . . .	21
Figure 2.7	Examples of basic arbiters, where R is request, P is priority, C is a carry, G is grant, and Any-G = $ (G_0 \dots G_{n-1})$ . . . . .	22
Figure 2.8	Structure of distributed cache for many multi-/many-core processors. . . . .	24
Figure 2.9	Projected delay issues of global regular wire compared to gate delay with technology scaling [7]. . . . .	27
Figure 2.10	Structure of the main three types of the transmissions lines: (a) microstrip line (MSL), (b) differential line or coplanar strips (CPS), and (c) coplanar waveguide (CPW). . . . .	33

Figure 2.11	Zenneck surface wave propagation decay which is significantly better than free space propagation [105]. . . . .	34
Figure 2.12	The corrugated surface. . . . .	35
Figure 2.13	Flat surface wave signal power loss proportional to frequency in the case of a dielectric material with relatively high loss tangent (Tan-Loss). These results obtained in collaboration with K.Tong in University Collage of London . . . . .	36
Figure 2.14	Comparison of forward transmission gain ( $S_{21}$ ) between Wireless [22] and SW interconnects [105]. . .	37
Figure 3.1	Integrated transceiver and integrated transducer (inverted quarter-wavelength monopole) stacked over the designed surface. . . . .	47
Figure 3.2	Surface wave interconnect communication channel with multi sub-channels where the master node transmit through the shared surface to slave node(s). . . . .	48
Figure 3.3	Experiment results showing power decay with distance for a range of frequencies [35]. . . . .	51
Figure 3.4	Measurement versus calculated voltage gain comparison for different frequencies. . . . .	51
Figure 3.5	Bit-Error-Rate vs. SNR for 16-QAM modulation sub-channel (SC). . . . .	52
Figure 3.6	The surface-wave experiment set-up. . . . .	53
Figure 3.7	The surface-wave experiment equipment alignment and set-up. . . . .	54
Figure 3.8	$S_{21}$ measurements results for the corrugated surface, flat surface, free-space, and the case where transducers connected face-to-face for wide range of frequencies. . . . .	55
Figure 3.9	$S_{21}$ measurements results for the flat surface shows signal attenuation with distances for wide range of frequencies. . . . .	55
Figure 4.1	Example showing that inserting two SWI channels in the proposed hybrid wire-SWI multilayer-network increases the overall NoC bisection bandwidth: (a) conventional on-chip network layer with 4-ary 2-mesh topology; (b) connections of both layers, metal wire and SWI. . . . .	60
Figure 4.2	Closing the gap towards small world phenomena in a $10 \times 10$ NoC as the number of nodes with transmission capability via Zenneck surface-wave interconnect (SWI) ( $N_m$ ) increased. . . . .	61

Figure 4.3	Example of 4 master placement in a $6 \times 4$ NoC based on simple simulated- annealing optimization algorithm with target to minimize the average-hop-count from all slaves to the nearest master. . . . .	62
Figure 4.4	An illustration example showing the master node routing decision, which is either forwarding the traffic via SWI or continue via regular Mesh. . . . .	64
Figure 4.5	Energy dissipation per bit according to on-chip communication distance for buffered wire [22] and SWI. . . . .	65
Figure 4.6	DWA implementation, where CSR provide the round robin functionality and store the weight code. . . . .	66
Figure 4.7	The conducted system-level evaluation flowchart shows the methodology and tools used to obtain the results. . . . .	67
Figure 4.8	$6 \times 4$ Network average delay verses PIR for W-SWI and baseline architecture. . . . .	68
Figure 4.9	$6 \times 4$ Network throughput verses PIR for W-SWI and baseline architecture. . . . .	69
Figure 4.10	Comparison between W-SWI with DWA algorithm and with basic Round-robin (RR) Communication power saving ratio to the Mesh for different network size and traffic scenarios. . . . .	71
Figure 5.1	(a) The non-trivial 1-to-M traffic percentage according to the simulation of a range of chip-multiprocessor (CMP) benchmark applications (from PARSEC and SPLASH2) with modified, exclusive, shared and invalid (MESI) cache coherence protocol; (b) our $6 \times 6$ regular mesh NoC simulations with random traffic plus random traffic with a small percentage of multicast or broadcast (5%). The introduction of multicast or broadcast leads to severe deterioration in performance in terms of latency and saturation packet injection rate (PIR). . . . .	76
Figure 5.2	Demonstration of different 1-to-M routing schemes: (a) path-based delivers packets sequentially in a worm-like route 5.2a, (b) Tree-based delivers the packets in a tree-like route [16, 15]. . . . .	78
Figure 5.3	Illustration example of multicast dependency that causes deadlock in tree-based multicast routing [156]. . . . .	78
Figure 5.4	Optical-based interconnect architecture examples that support on-chip multicast. . . . .	80

Figure 5.5	Examples of multicast clustering in WiNoC-based interconnects architectures. . . . .	80
Figure 5.6	Examples of some RF-I multicast architectures [18, 19]. . . . .	82
Figure 5.7	W-SWI improved tree-based with low latency packet delivery where branching possible only at the SWI master nodes . . . . .	84
Figure 5.8	Demonstration of the deadlock problem created by the multicast dependency. . . . .	85
Figure 5.9	The proposed MAB-check unit to mask the requests unless all the requested multicast group are free. . . . .	86
Figure 5.10	Design of the proposed global multi-resource arbiter (GMA) for SWI channels: stage (1) request masking; stages (2-4) achieve legal match with lonely output allocator [51]; stage (5) generates the grant signals for a fixed period. The figure also shows an example of GMA stages 1-4 with four masters and two VCs. Master ( $M_4$ ) related logic is not drawn, for simplicity, but it is currently allocating some of the slaves requested by $M_2$ . . . . .	89
Figure 5.11	Demonstration control signals of SWI communication protocol exchanged between master, slaves and the GMA. . . . .	91
Figure 5.12	Average delay results of $6 \times 4$ NoC with the following: (a) comparison of Mesh and W-SWI-C under random traffic with 10% broadcast; (b) W-SWI-C with different allocation techniques; (c) W-SWI-C with different SWI master number. Note that W-SWI: $N_m$ : $VCN_v$ :ArbP refers to W-SWI with $N_m$ number of masters, $N_v$ number of VC and P number of grant cycles. Hold is where GMA grants a master until all its current data flow is transmitted. . . . .	93
Figure 5.13	Communication power saving ratio of the W-SWI over the Mesh for different network sizes, types of synthetic traffic and broadcast/multicast ratios. . . . .	95
Figure 5.14	Packet delay distribution comparison of W-SWI and Mesh with software multicasting for a $6 \times 4$ NoC with 5% broadcast. . . . .	96
Figure 5.15	Average delay comparison of NoC ( $6 \times 4$ ) between W-SWI and VCT and Mesh under SNN benchmarks for different NoC size. . . . .	98

Figure 6.1	Example shows traffic flows from four masters multiplexed in SWI where (a) W-SWI-C architecture with Hold allocation mechanism (b) W-SWI-C architecture with fixed period (one cycle) alternation and two VSWI, and (c) W-SWI-D architecture (decentralized arbitration and stretched multicast), where M is a master, S is a slave and T is a time slot. . . . .	103
Figure 6.2	Illustration of router micro-architecture with ID-Tagging based flow control and VC flow control. . . . .	105
Figure 6.3	Simulation flow to obtain the results. . . . .	109
Figure 6.4	Comparison between the average delay of W-SWI-C and W-SWI-D under uniform synthetic traffic with 10% multicast ratio for different VC number and NoC size. . . . .	111
Figure 6.5	Comparison between the average delay of W-SWI-C and W-SWI-D under uniform synthetic traffic with 5% and 10% broadcast ratios. . . . .	112
Figure 6.6	Comparison between the average delay and energy improvements of W-SWI-C and W-SWI-D over Mesh under real applications benchmarks from PARSEC [150] and SPLASH2 [151] for $10 \times 8$ NoC. . . . .	115
Figure A.1	(a) Demonstration of router ports with three VCs; (b) The added control signals between tiles to simulate three VCs in this example. . . . .	123

## LIST OF TABLES

---

Table 2.1	Summary of reported key features for implementation of integrated optical interconnects. .	28
Table 2.2	Examples of demonstrated integrated wireless communication systems along with their key features for a single link. . . . .	30
Table 2.3	Examples of reported implementations of integrated transmission lines along with their key features for a single link. . . . .	32
Table 2.4	Summary comparison of key features in current and emerging on-chip interconnects. . . . .	38
Table 2.5	Comparison of various NoC models and simulators.	43
Table 4.1	The adopted parameters of the NoC based multi-processor chip ( consists of $6 \times 4$ tiles). . . . .	63

Table 4.2	Scaling of number of SWI channels as the NoC scale based on the assumption that the complementary metal-oxide-semiconductor (CMOS) technology has scaled. . . . .	67
Table 4.3	W-SWI PIR and throughput improvement over Mesh at the edge of network saturation where latency reach double the zero-load-latency (ZLL). . . . .	70
Table 4.4	W-SWI average delay and throughput improvement over Mesh for a range of applications . . .	70
Table 4.5	Area overhead consideration for the proposed Architecture comparing to related Architectures. Noticed that some component are needed only in the Master nodes . . . . .	72
Table 5.1	Average delay and PIR at the edge of NoC saturation comparison of the W-SWI and VCT. . . . .	97
Table 5.2	Area overhead evaluation for W-SWI-C, proposed W-SWI-D and VCT-512 [15] over baseline architecture (Mesh). . . . .	99
Table 6.1	Notations used in this chapter. . . . .	102
Table 6.2	Comparison of highlighted features of the centralized and decentralized approaches for the proposed architecture. . . . .	106
Table 6.3	Results improvements over Baseline architecture (Mesh) comparison between W-SWI-C and the proposed W-SWI-D with 10% multicast ratio. . .	114
Table 6.4	Area overhead evaluation for W-SWI-C, proposed W-SWI-D and VCT-512 [15] over baseline architecture (Mesh). . . . .	116

## LIST OF ALGORITHMS

---

4.1	Distance based weighted random arbitration algorithm(DWA). . . . .	65
5.1	procedure of centralized arbitration and allocation using global-multiresources-arbiter (GMA). . . . .	87
6.1	Master interfaces procedure to communicate via SWI. . . . .	107
6.2	slave interfaces procedure to communicate via SWI. . . . .	108

- A.1 Algorithm of the generation of one-to-many (1-to-M) or one-to-all (1-to-all) traffics and the way of injected them based on the architecture that handel the multicast-traffic; Where N is the number of nodes and Flit.M is the multicast flag bit. . . . . 124

## ACRONYMS

---

1-to-M	one-to-many
1-to-all	one-to-all
1-to-1	one-to-one
2D	two-dimensional
3D	three-dimensional
3D-ICs	Three-dimensional integrated circuits
BER	bit error rate
CDMA	code-division-multipliable-access
CMP	chip-multiprocessor
CMPs	chip-multiprocessors
CMOS	complementary metal-oxide-semiconductor
CPS	coplanar strip
CPW	coplanar waveguide
CSR	circular-shift-register
DOR	dimension-ordered routing
DVFS	dynamic voltage and frequency scaling
DWA	Distance-based weighted-round-robin arbitration
EM	electromagnetic
FIFO	First-In-First-Out
FDMA	frequency-division-multipliable-access
GPU	graphics processing unit
GMA	global-multiresources-arbiter



ID	identity
IC	Integrated circuit
IP	intellectual property
IPs	intellectual properties
LLC	last-level-cache
LSB	least significant bit
LOA	lonely output allocator
MAB	multicast-address-bits
MESI	modified, exclusive, shared and invalid
MPSoCs	multiprocessor systems-on-chip
MSL	microstrip line
NI	network interface
NIs	network interfaces
NoC	network-on-chip
NoCs	networks-on-chip
OOK	On-off keying
ONoC	optical network-on-chip
ONoCs	optical networks-on-chip
PE	processing element
PEs	processing elements
PIR	packet injection rate
QAM	quadrature-amplitude-modulation
QoS	quality of service
RF	radio frequency
RF-I	RF-interconnect with transmission-line
RF-Is	RF-interconnects with transmission-lines
RR	round robin
Rx	receiver
SWIs	Zenneck surface-wave interconnects

SC	sub-channel
SCs	sub-channels
SoC	system-on-chip
SoCs	systems-on-chip
SNR	signal-to-noise-ratio
SNN	spiking-neural-network
SNNs	spiking-neural-networks
SCC	single-chip cloud computing
SWI	Zenneck surface-wave interconnect
SW	Zenneck surface-wave
TSVs	through-silicon-vias
TSV	through-silicon-via
TL	transmission-line
TLs	transmission-lines
Tx/Rx	transceiver
VCs	virtual channels
VC	virtual channel
VLSI	very large scale integration
VCT	virtual-circuit-tree
VSWI	virtual-channel for surface-wave-interconnect
VSWIs	virtual-channels for surface-wave-interconnect
WDM	wavelength-division-multiplexing
W-SWI	wire and surface-wave interconnects
W-SWI-C	W-SWI centralized
W-SWI-D	W-SWI decentralized
WiNoC	wireless network-on-chip
WiNoCs	wireless NoCs
ZLL	zero-load-latency

Part I

Thesis Chapters

## INTRODUCTION

---

### 1.1 MOTIVATION

Market demand for ever more complex and sophisticated digital electronic systems has led to the exponential scaling of integrated circuit technology processes, as predicted by Moore's law. In addition, the scaling cost and limited bandwidth of off-chip communication have motivated the design option of integrating more on-chip system components to keep most communication on-chip [1]. This causes an intensification of current and future system-on-chip (SoC) in terms of the number of integrated intellectual property (IP), which includes processing elements, memories and application-specific intellectual properties (IPs). Consequently, the volumes and numbers of communications have also increased dramatically. As a result, a bus-based interconnect architectures, which have been managing the on-chip communication since the introduction of SoC, could no longer satisfy the communication requirements.

Therefore, the research community [2, 3, 4] and industrial organizations [5, 6] have adopted the network-on-chip (NoC) as a scalable underlying communication structure over the regular bus-based interconnect architectures. Although networks-on-chip (NoCs) offer system designers scalable and adaptable on-chip communication and a plug-and-play design approach, the fundamental physical channels, where the signal is transmitted by charging and discharging the whole wire, remain the same. Therefore, as capabilities of the physical-foundation of the on-chip interconnects are stretched to their limits with each technology scaling, a growing burden is placed on the wire-based NoC.

This is due to the fact that the downscaling of the complementary metal-oxide-semiconductor (CMOS) targets gate nodes rather than wires in terms of performance and power consumption. Therefore, as we enter the sub-micron zone, on-chip communication becomes a performance bottleneck and dominates power consumption. Thus, regular metal-based NoCs struggle to match this scalability, especially for global communication, in terms of latency and energy (J/bit) [7, 8]. Moreover, the cross-section of the wires is decreasing [7] and thus its resistivity is increasing, which causes higher power dissipation. Therefore, the issue is not only that metal wire does not scale enough to match future interconnect requirements, but also that its performance is projected to get worse in terms of power and latency.

Moreover, in addition to these wire issues, given the nature of multi-hop NoC communication, providing efficient across-the-chip communication is also becoming a serious challenge. Some studies have proposed 3D-integration to ease global communication issues by reducing the average NoC hop-count. However, although promising, this technology faces various technical challenges such as process control requirements, wafer thinning, low through-silicon-via (TSV) capacitance and design challenges [9, 7, 10].

These intra-chip communication challenges are especially pertinent for chip-multiprocessors (CMPs), which were introduced to provide near-linear performance improvements when complexity increases while maintaining lower power and frequency budgets [11]. Therefore, the International technology roadmap for semiconductors (ITRS) predicts that we will reach the many-core era in only few years where hundreds of cores are integrated, as shown in Fig. 1.1. This will be realized in the very near future since some graphics processing unit (GPU) chips has been already fabricated with thousand cores [12]. The increase in numbers of cores, but not their complexity, along with the projected wire issues, will increase the proportion of power budget accounted for communication rather than computation. Also, these factors will make performance increasingly determined by inter-core communication rather than core-complexity. For instance, the NoC power budget for  $8 \times 10$  CMP TeraFLOPS, is reported to be around 35% [6]. Consequently, many-core systems design has shifted from computational-centric to communication-centric [13].

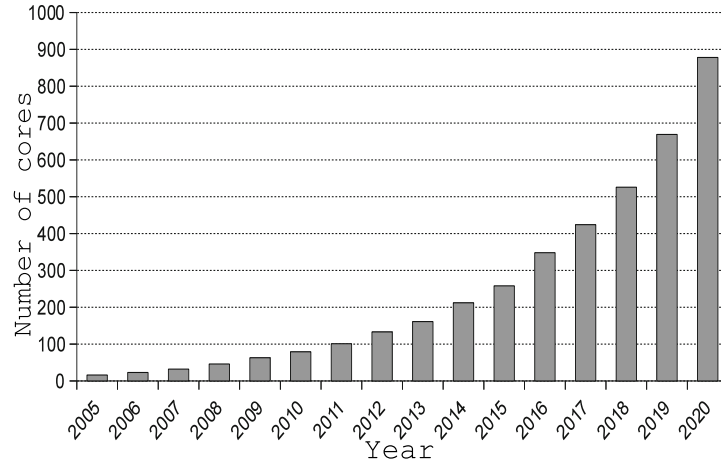


Figure 1.1: Processing element number is projected to scale exponentially according to ITRS SoC [14].

On the other hand, multi/many-core performance and power consumption depends not only on the NoC, but also on cache coherence protocols. Cache coherence protocols generate a range of multicast (one-to-many (1-to-M)) or broadcast (one-to-all (1-to-all)) communication patterns [15, 16]. These types of traffic are projected to scale in terms

of number of destinations, burstiness, and spatial distribution as the number of cores scale up [17]. For instance, broadcast-based cache coherence protocols produce a relatively high broadcast ratio over total packet injection rate (PIR) of up to 52.4% [16, 15], as shown in Fig. 1.2a. This could be catastrophic for global coherence and NoC performance unless the on-chip interconnect fabric supports 1-to-M communication. Fig. 1.2b shows the severe effect of broadcast traffic scaling on the NoC performance. Therefore, there is a need to overcome these constraints and improve performance by using new interconnect architectures that support multicast. Relevant NoC studies, at most, aim to achieve 1-to-M latency and energy levels close to wire-latency and wire-energy [16, 15]. Again, this will not be sufficient in the near future given the projected issues with regular metal-based NoCs, since these interconnect fabrics struggle to match the required scalability in terms of latency and energy (J/bit) [7, 10].

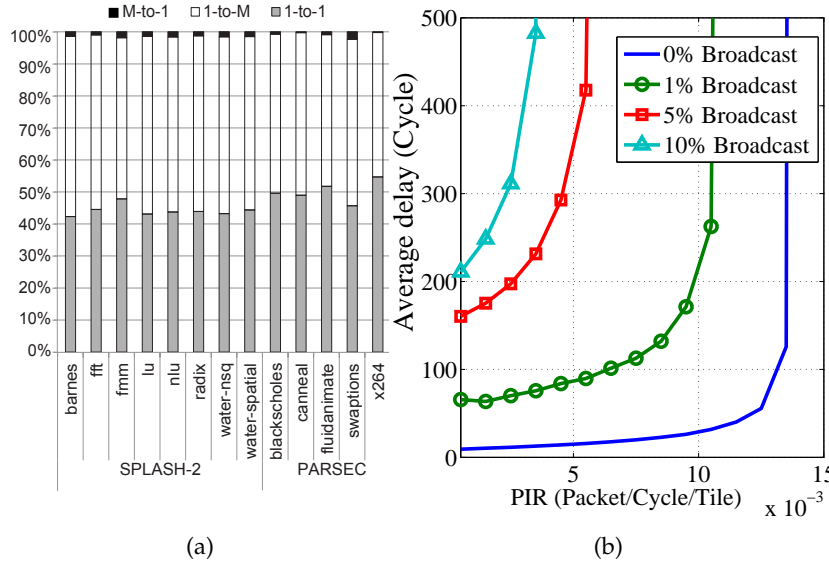


Figure 1.2: (a) 1-to-M, M-to-1 and 1-to-1 traffic percentage in different CMP real benchmark when a broadcast-based cache coherence protocol (token coherence) is used [16], (b) NoC simulation shows an increase in average delay and fast network saturation when broadcast treated as unicast in a  $6 \times 4$  NoC.

Thus, the challenges facing global and multicast communication have inspired many researchers to look for alternative or supplementary types of interconnect. Such emerging cutting-edge interconnects include the RF-interconnect with transmission-line (RF-I) [18, 19, 20, 21], wireless network-on-chip (WiNoC) [22, 23, 9, 24, 25, 26] and optical network-on-chip (ONoC) [27, 28, 29, 30]. However, these types of interconnect themselves face various challenges, such as their complexity, power consumption and/or area overheads.

This thesis opens a new research directions in the development of new and promising on-chip communication technologies. This

study by proposing the Zenneck surface-wave interconnect (SWI) for intra-chip interconnects has already inspired several research teams to further investigate this promising technology [31, 32, 33]. This technology involves an electromagnetic wave that propagates, and is guided, through an interface between the surfaces of different media. It has many remarkable features compared to state-of-the-art emerging interconnects in terms of low-power dissipation, high signal fan-out and close to the speed of light cross-the-chip propagation [34, 35, 36, 37, 38]. In addition, this thesis proposes innovative architectures and technologies that maximize the utilization of the SWI and resolve their system-level challenges. As a result, the original proposed architecture demonstrates a promising and radical solution to meet future NoC-based multi/many-core processors necessities such as those concerning global and multicast communication.

## 1.2 STATEMENT OF ORIGINALITY

The major contributions of this thesis can be summarized as follows:

- A comprehensive view is presented of current knowledge of the merits and drawbacks of emerging interconnects such as WiNoC, RF-I, ONoC, and SWI. Subsequently, this survey could be a ground for any study inspired to utilize these emerging interconnects advantages and address their challenges. In addition, a system-level comparison of these promising types of interconnects is provided. This comparison especially highlights the future communication functionality requirements to the under-layer physical and technological capabilities of these fabrics [38].
- The challenges, performance, reliability, implementation, and design considerations for the SWI are studied. This has led to a quantification of the effectiveness of this promising interconnect. Moreover, an analytical model for SWI links power dissipation is developed based on previous experimental results. The calculated metrics are then used for subsequent system-level evaluations [35, 34].
- Experiments on different types of designed waveguide surface have been conducted to build the ground for understanding of SWI features [39].
- A wire and surface-wave interconnects (W-SWI) interconnect architecture is proposed. This NoC architecture aims to utilize the SW features to allow future on-chip communication scalability. In addition, these efficient short-cut links are placed so that the resulting network topology reduces the intra-core average hop count. This two-layer NoC architecture has considerable advan-

tages in terms of connectivity, since it employs the merits of both mesh and configurable bus topologies [35, 34].

- The hybrid W-SWI architecture is improved to mitigate global communication issues. This is achieved by developing a Distance-based weighted-round-robin arbitration (DWA) arbitration technique. This simple and efficient technique intelligently prioritizes surface-wave channels over metal wires, without saturating these channels, in the case of global communication. The proposed DWA is synthesised and tested so that measurements can be used in a system-level simulator. Evaluations are conducted using this cycle-accurate simulator, which show improvements in average delay (34%), throughput (35%), and power consumption (12 to 23%). In addition, these improvements are achieved with negligible cost [35].
- A tree-based multicast routing scheme is developed that exploits the W-SWI architecture for 1-to-M traffic handling. This routing scheme enabled by SWI is superior to other state-of-the-art multicast/broadcast routing schemes in terms of latency and power consumption, since it forks the packet only at the destination routers. In addition, it requires only relatively minor alteration to the baseline router micro-architecture [36, 40]
- A W-SWI centralized (W-SWI-C) is proposed for 1-to-M traffic that efficiently addresses multicast traffic contention issues and maximizes SWI utilization. In particular, a novel global-multiresources-arbiter (GMA) is designed and a detailed design rationale, hardware realization and schematic are presented. This centralized arbitration and allocation unit has the ability to allow the concurrent utilization of many resources with relatively low circuit complexity and delay [36, 40].
- The W-SWI-C is rigorously evaluated using a cycle-accurate simulator for both synthetic traffic and real application benchmarks. Moreover, the proposed GMA is synthesised and tested. This proposed architecture is found to surpass previous related work by achieving improvements in average delay ( $\sim 22\times$ ), power consumption ( $\sim 2 - 10\times$ ), and attaining a reliable quality of service. Moreover, the additional hardware cost of the W-SWI is found to be relatively insignificant [36, 40].
- A W-SWI decentralized (W-SWI-D) architecture is introduced which deploys novel techniques such as stretch-multicast and ID-tagging flow control. These efficient and flexible decentralized arbitration schemes maximize the hybrid architecture's utilization and contention handling for 1-to-M traffic. These techniques have been mathematically proven to be deadlock-free [37, 40].



- The proposed W-SWI-C is compared and evaluated thoroughly against W-SWI-D in terms of latency, power consumption and area overheads. The W-SWI-D has been found to perform better when traffic levels are relatively low, but worse when the load reaches the NoC saturation points and case of broadcast. The evaluation also demonstrates that W-SWI-C has less area overhead compared to W-SWI-D, but both have relatively negligible area overheads compared to state-of-the-art on-chip multicast interconnect architectures [37, 40]

### 1.3 THESIS ORGANIZATION

This thesis is organized into seven chapters, as shown in Fig. 1.3. Two main requirements of future many-cores are tackled, which are global and multicast communication, using the proposed hybrid wire-SWI architecture. The Surface-wave interconnects is covered in Chapter 3, while Chapter 4 introduces the proposed hybrid architecture and addresses the on-chip global communication requirements. On the other hand, Chapters 5-6 cover original techniques based on the proposed architecture to handle multicast communication challenges.

Chapter 1 "Introduction": introduces the motivations, objectives, contributions and structure of this thesis.

Chapter 2 "Background and Literature Review": provides background information and summarizes the literature on topics relevant to this thesis. In addition, the emerging on-chip interconnects included in recent literature are reviewed and discussed in detail.

Chapter 3 "Surface-wave On-chip Interconnects": proposes the novel SWI, discusses its merits and challenges, and develops an analytical model of power consumption and performance for the SWI.

Chapter 4 "Hybrid Wire and Surface-wave Architecture for On-chip Global Communications": proposes the Hybrid wire and surface-wave interconnects architecture W-SWI and discusses the utilization and efficiency of this proposed architecture for global/semi-global on-chip communication.

Chapter 5 "Wire and Surface-wave Architecture with Centralized Control for Multicast" proposes an original deadlock-free centralized arbitration and allocation technique that exploits the merits of the W-SWI architecture for 1-to-M communication. Moreover, it presents a comprehensive evaluation of the proposed architecture and techniques compared to state-of-the-art architectures running real benchmarks.

Chapter 6 "Wire and Surface-wave Architecture with Decentralized Control for Multicast" further explores multicast contention issues by developing new decentralized flexible schemes. A comparative evaluation of centralized decentralized approaches is also presented.

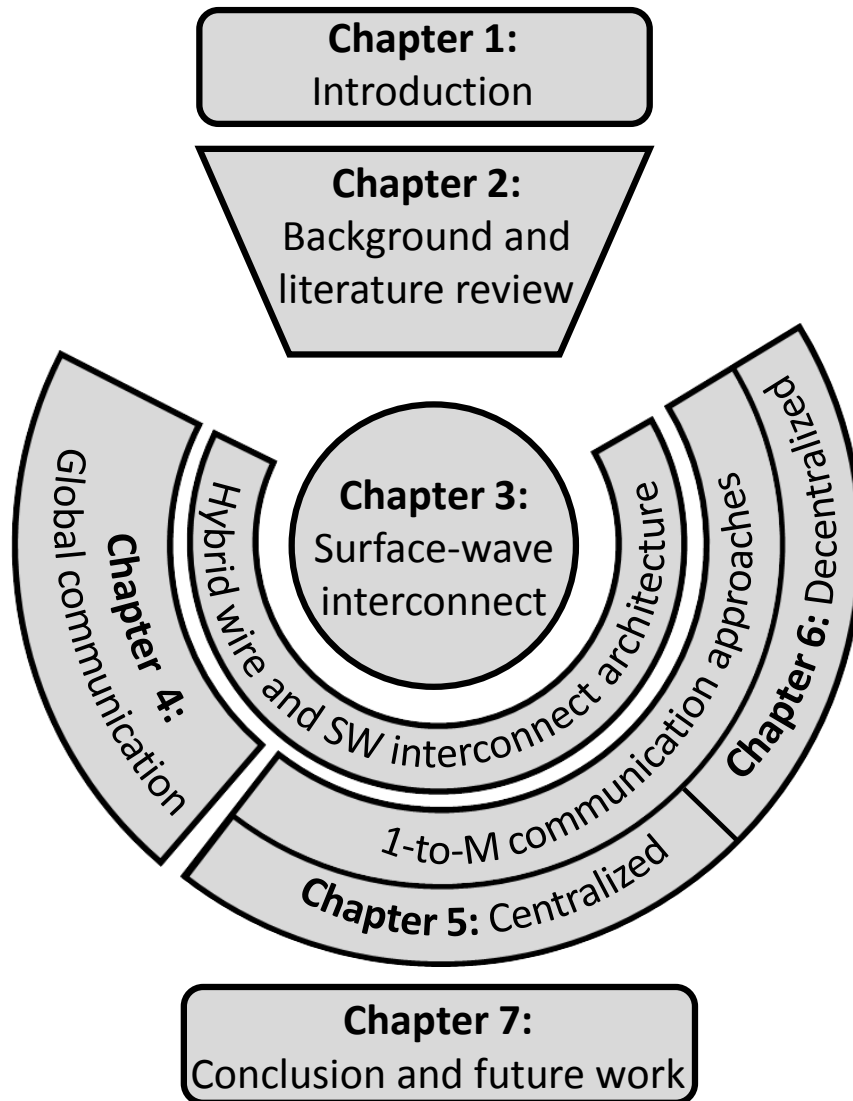


Figure 1.3: Thesis organization.

Chapter 7 "Conclusions and Future Work" summarizes the conclusions of the study and discusses the implications of the presented research and draws the horizon for potential future work.

## BACKGROUND AND LITERATURE REVIEW

---

### 2.1 INTRODUCTION

Growing demand for complex digital systems has caused a dramatic increase in the number of integrated intellectual properties (IPs) in current and future system-on-chip (SoC). For instance, in the last decade, in order to maintain balance between further performance improvements and available power and frequency budgets, the use of small many-cores is preferred over single or small numbers of complex cores [11]. As a result, the number of integrated IPs/cores inside a single SoC has increased dramatically.

On the other hand, volumes of inter-core or IP communication are increasing significantly. In contrast, technology scaling enables the realization of complex, yet with economic implementation, of many computational or control components. Therefore, on-chip communications are becoming a performance bottleneck for future systems-on-chip (SoCs), including chip-multiprocessors (CMPs), due to the growing burden caused by scalability requirements. As a result, a communication-centric design has become vital for current and future SoCs [41, 13].

Consequently, the network-on-chip (NoC) has been proposed for over a decade now as a way to tackle the challenges and requirements of on-chip communication. Since then, rapid advances in NoC design have been achieved by the research community [2, 3, 4, 42, 43, 44] and industry [5, 6, 45, 46, 47].

This chapter reviews the main concepts of NoC including architectural components, topologies, routing, switching and flow control. Moreover, cutting-edge current and emerging types of interconnects are discussed along with recent ground-breaking advances in these fields by academic and/or industrial entities.

### 2.2 BACKGROUND

#### 2.2.1 *Networks-on-chip*

On-chip communication since the introduction of the SoC had been limited to either bus architecture, point-to-point interconnect, or a mix of both until the last decade. In the former, the bus consists of single shared wires (a physical channel) that is accessible by a set of logical components via their logical channels. Even though this architecture has relatively low area overheads and high fan-out capability, it suffers

from performance and power limitations [48]. This is still true even when the bus architecture evolved into a hierarchical bus of which segments of the bus are connected via bridges that could buffer the data, such as the AMBA [49].

The second architectural approach is the point-to-point wire interconnects. This is the simplest type of interconnect, which consists of dedicated physical channels that link each two components in the system. However, the number of channels grow rapidly as we increase the numbers of IPs or cores ( $N$ ) that need to be connected ( $O(N(N-1))$ ). In addition to the cost of many such physical channels, the routing of these wires inside the chip could become a nightmare.

The NoC has emerged as the underlying on-chip communication structure, which offers the best cost/performance trade-off that meets scalability requirements. These networks-on-chip (NoCs) were inspired by data communication networks, and such an architecture consists of a network constructed from multiple point-to-point data channels (links) interconnected by routers. The routers are connected to a set of distributed IPs/cores and communication among these usually utilizes a packet-switching method where messages are divided into suitably-sized blocks, called packets. This section briefly introduces the main concepts of the NoC.

#### 2.2.1.1 NoC Topology

Network topology is the arrangement or pattern in which the network nodes are connected using physical channels. For the SoC, these nodes could refer to any intellectual property (IP) component such as processing elements (PEs) and memories in the case of a direct network or routers and switches in the case of an indirect network [3]. However, since indirect networks provide more efficient connectivity with large numbers of IPs, mostly the NoC are indirect networks and the nodes are referring to routers. The first step in designing a network is to determine its topology that should balance connectivity requirements and the resources available. In other words, the best trade-off between performance and cost should be found. The second step is to ensure that the resulting topology connects all of the nodes.

Fig. 2.1 illustrates a few common on-chip network topologies. Although the bus shown in Fig. 2.1a was until the last decade the most common topology, scalability requirements have necessitated the search for alternatives. Fig. 2.1b presents a ring topology that, although slightly better than the bus, still suffers from fast network saturation. The butterfly topology shown in Fig. 2.1c has a low network diameter and known packet delivery delay. However, it does not offer path diversity unless a further layer of routers is added and it cannot be implemented without relatively long wires. The same drawbacks exist in crossbar networks in terms of a lack of path diversity and the need for long wires, as shown in Fig. 2.1d. Moreover, even though it

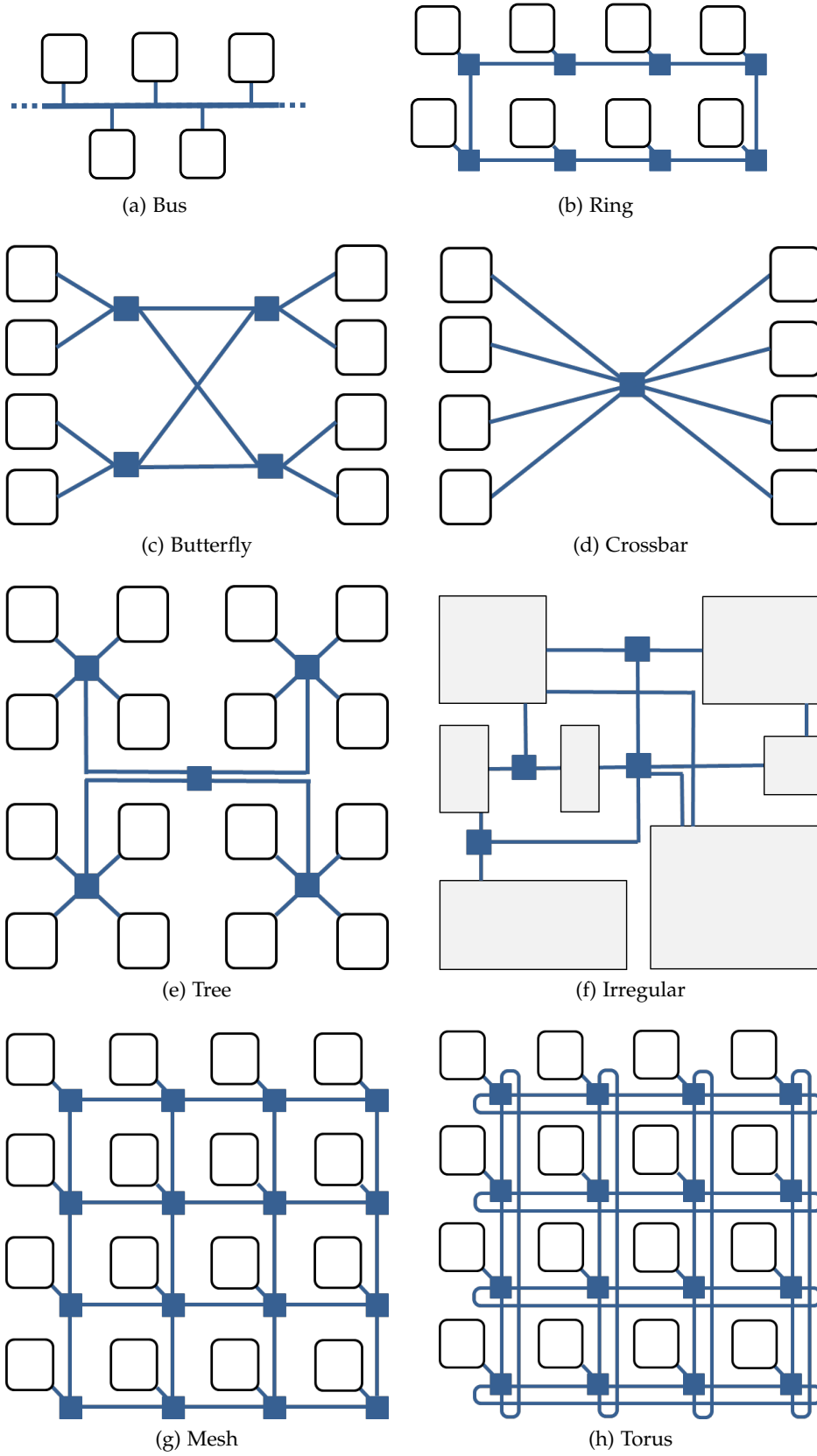


Figure 2.1: Examples of on-chip communication network topologies.

involves the attractive one-hop communication, the cost of the crossbar increases exponentially as the number of nodes are increased. The tree topology shown in Fig. 2.1e has low zero-load-latency, but networks are also saturated very quickly as the packet injection rate (PIR) increases since the root becomes a bottleneck.

The torus and mesh topologies demonstrated in Fig. 2.1h and Fig. 2.1g, respectively are the most common topology for general purpose NoCs. This is due to their many desirable features such as physical arrangement suitable for chip floor planning. Moreover, the uniform physical arrangement benefits applications with high communication locality characteristic where communication with neighbours is more common than with the rest of the network. In addition, these topologies have characteristics of path diversity, load balancing, and high throughput. The mesh has shorter wires than torus especially if the latter is unfolded. Moreover, although the torus has a lower average hop count, both the mesh and torus have higher average hop counts than the butterfly, crossbar and tree topologies.

On the other hand, there are some irregular application-specific network topologies where the designer tailors the network to the application's communication requirements, as in Fig. 2.1f. However, such networks have very low adaptivity to changes in the injected application load. In addition, they also have the drawback of long wires and/or non-uniform and high switches degrees. As a result, these network topologies are very limited. On the other hand, some designers have tried to combine more than one topology to utilize benefits of both topologies, but the resulted topology performance should justify the extra cost involved [50].

#### 2.2.1.2 NoC Component

The NoC may include different devices based on architecture and design requirements. This section briefly describes the most common NoC components. Fig. 2.2 shows a regular NoC with mesh topology and how its components such as routers, links, and network interfaces (NIs) are connected. In addition, a block diagram of the micro-architecture is given of a typical router and network interface (NI).

The router is responsible for directing the packets towards their destination. In addition, the router applies the routing algorithm and performs the flow control, as will be discussed in Sections 2.2.4 and 2.2.3. The router's building blocks can be classified as either data path or control plane as shown in Fig. 2.2. The data path includes all the units that store and pass packets such as input/output buffers and crossbar switches. On the other hand, control plane units manage and coordinate the movements of packets in the data path units.

Once the packet arrives at the input port, it is stored in the FIFO buffer. Then the routing information is extracted and sent to the routing unit that determines the possible output port(s). In cases

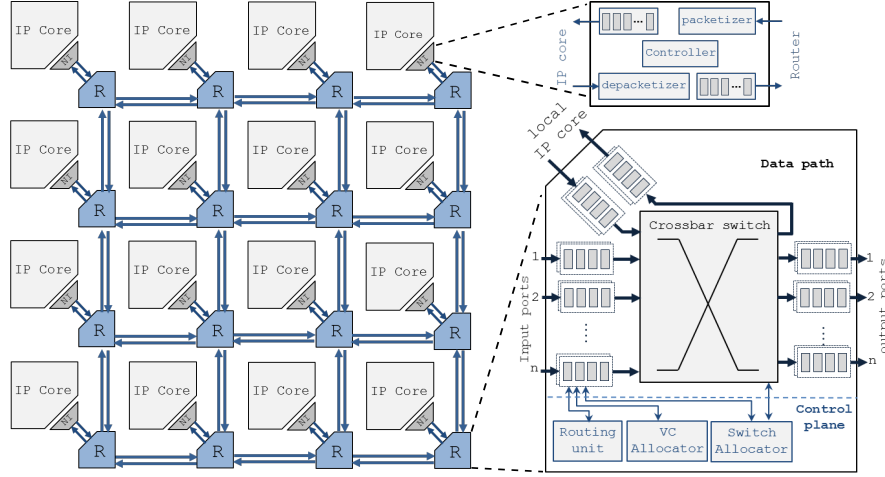


Figure 2.2: main components of NoC such as router(R) and network interface (NI) common micro-architecture.

where the router design includes multiple virtual channels (VCs) per port, the virtual channel (VC) allocator and arbiter reserve the output VC for the specific input VC. After that the switch allocator reserves switch time slots, since it is only allowed to link at most one input port to at most one output port. As a result, the data passes to the output port buffer where it must wait until it is transmitted via the physical channel. These pipeline steps might be different and/or completed simultaneously to improve crossing latency depending on the router design. Moreover, these pipelined steps are either conducted per flit or packet. For instance, routing and VC allocation are conducted per packet only. This reduces the dynamic power and crossing latency [51]. More information about the flow control of packets and flits will be discussed in Section 2.2.3.

Network interfaces are much simpler than routers. Their task is simply to link the local IP cores with the routers. This task requires the packetization and depacketization of the payload message. Also, it may include the temporary storage of the packets to and from the local IP cores.

### 2.2.2 Deadlock, Livelock and Starvation

A packet might not reach its destination, or may fail to progress even though there is no failure in the network. This is due to problems such as deadlock, livelock and starvation. Generally, these problems are caused because of finite network resources. Therefore, careful consideration is required when designing the techniques used to forward packets from their source to destination in order to avoid or recover from these problems.



Channel starvation can result from a request for a network resources never being granted because the arbitration always or frequently grants other requests. For instance, packet starvation will exist if a packet keeps requesting a crossbar switch time slot, which is always granted to other ports. This issue can easily be handled by adopting a fair arbitration and allocation mechanism.

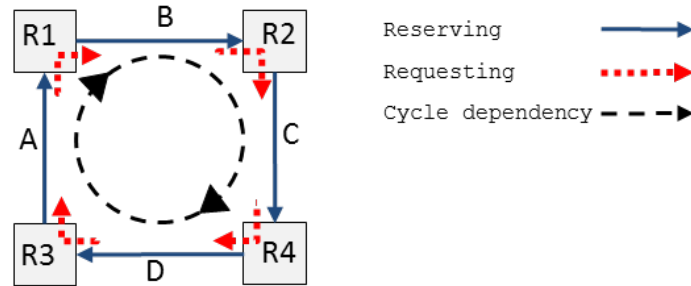


Figure 2.3: Example scenario of a deadlock resulting from cyclic channel dependency.

Deadlock results from the cyclic dependency of packets reserving some network resources and requesting those of each other. For example, Fig. 2.3 shows how packets A, B, C, and D are requesting channels reserved by B, C, D, and A, respectively, which creates a cycle of dependency. This situation will prevent any of these packets from progressing unless one or more of them releases the channels it has already reserved. This is a very serious problem that might paralyse the whole network since it increases the number of blocked packets even if they are not part of the deadlock scenario [52]. This problem is usually solved either by deadlock avoidance or deadlock recovery. The former tries to remove the conditions that lead to the deadlock scenario, such as restricting some routing turns or removes packet dependency by implementing virtual channels, see sections 2.2.3.5 and 2.2.4.1. On the other hand, deadlock recovery tries to detect a deadlock occurrence and then recover from it.

The third problem, which is livelock, is usually associated with nonminimal adaptive routing, which discussed in Section 2.2.4. In this case, the packets are not stopped from progressing, yet they also do not reach their destinations. This is due to the packet continuously being misrouted in a cyclic manner due to the blocking or congestion of channels in the path toward the destination. The solution to such a problem includes restricting some paths, adopting minimal path routing, limiting the number of misroute, or using probabilistic avoidance [3, 51].

### 2.2.3 Flow Control

Flow control is the techniques to efficiently balance network resources such as buffers and physical channels among travelling messages. An effective flow control utilizes these resources and achieves higher bandwidth and lower latency. This can be viewed as resources allocation problem, so that these resources should be allocated more efficiently. Also, it can be considered as contention handling when two messages requesting the same resources [51].

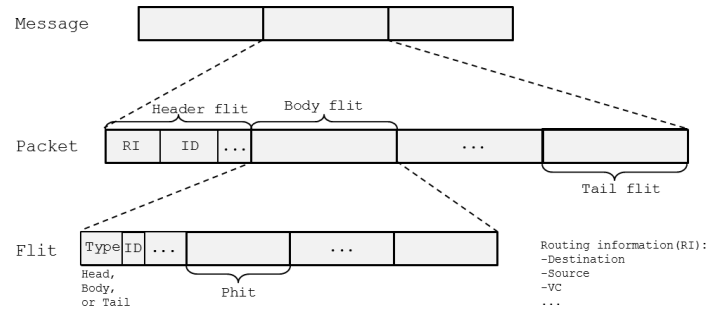


Figure 2.4: illustration of segmentation and the encapsulation process of message into packets, flits, and phits depending on the NoC resources and the flow control techniques.

Based on the resources available and the flow control techniques used, the data messages might need to be segmented into packets. This segmentation process is called packetization, which is achieved by the NIs. The data segment or payload is encapsulated by the packet header and tail that include information such as the destination, packet ID number, time stamp, and various status flags. For the same reason, the packet might also be divided into smaller units called flits. These are either header, tail, or payload flits and each also has its own header. A flit might also be divided into even smaller segments called phits. Fig. 2.4 shows the segmentation and structure of packets, flits and phits.

The following sections describe the main flow control techniques along with their advantages and disadvantages.

#### 2.2.3.1 Circuit Switching

Also known as bufferless flow control, this basically sets up the physical channel from the source to the destination using a routing header or probe that is transmitted before the data. After receiving an acknowledgement that a physical channel is reserved, the data is transmitted. This type of techniques is suitable for infrequent and long messages and where buffer avoidance is one of the design requirements.

### 2.2.3.2 *Store and Forward*

Also known as packet switching, this can buffer the whole packet in intermediate routers. In this way, there is no need to wait until all physical channels are free. Moreover, This also will improve physical links utilization by avoiding keeping segments of physical links idle and no longer needed until the packet reaches its destination. As a result, network efficiency in terms of latency and throughput is increased due to the network's ability to handle more packets simultaneously.

### 2.2.3.3 *Virtual Cut-Through*

In this technique, the router does not need to wait until the whole packet arrives. Instead, once the packet header that contains all the routing information has arrived, it starts to forward the packet. These pipelined packets improve further the effectiveness of the network. However, in highly congested networks the packets will be delayed for longer in each node, and thus the network will behave like a store-and-forward flow control.

### 2.2.3.4 *Wormhole*

Due to increased message size, the buffer size might be too small to hold the whole packet. Therefore, allocation can be accomplished on a flit basis and the packet might be extended over a few routers. The benefits of this technique are the low hardware resources needed, such as buffer size and physical channels, as well as its efficiency. Routing and port allocation are required only for the header flit, which is routed ahead of the rest of the packet. Then the other flits follow the same path as the header flit where they only need to reserve the switch time slot, as shown in Section 2.2.4. When all flits have been forwarded, the tail flit will release the reserved channels. Even though the wormhole is considered a very cost-effective technique, it makes the network more vulnerable to deadlock and contention issues since each packet might reserve several channels simultaneously. Therefore, careful consideration should be given to routing and contention handling techniques.

### 2.2.3.5 *Virtual Channels (VCs)*

Router buffers usually operate as FIFO queues. Therefore, if the packet is blocked because of congestion, for example, all the packets behind it will be blocked as well. Alternatively, the input and output buffers can be divided into separate FIFO queues that can be allocated separately. These are widely known as VCs, which were first proposed to avoid deadlock in wormhole flow control [51, 3]. This flow control technique allows the decoupling of physical channels and buffer allocation, which increases the flexibility of sharing network resources. However,

the flits in each VC lane have to compete for physical channels. This requires the addition of a VC allocation unit, as discussed in Section. 2.2.1.2. This unit might add an extra pipeline stage for the router crossing.

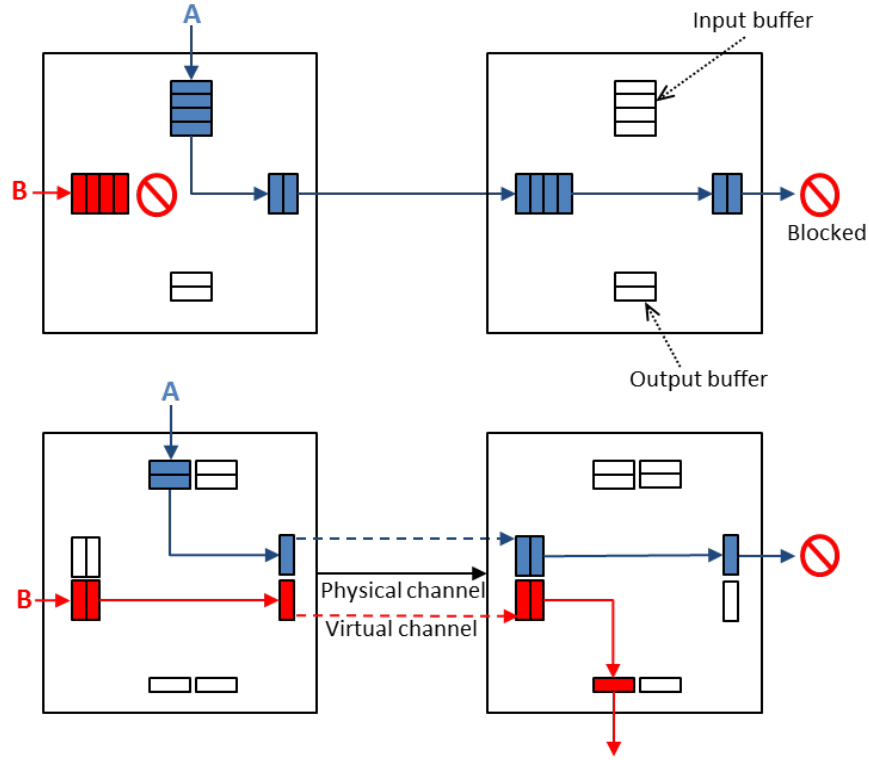


Figure 2.5: Example scenario showing the benefit of using VCs on mitigating blocking channels issues, even though the buffer size remain the same.

Each VC buffer either can have a static size or dynamically allocated size from a main buffer. Moreover, buffer size does not need to be increased when it is replaced by several queues. According to Dally [53], network will achieve high throughput and latency improvements by increasing the number of VCs even when the buffer size is kept constant. However, there is a limitation to the increase in the number of VCs where network performance starts to decay due to the latency in allocation and arbitration processes for high numbers of VC lanes. On the other hand, the VC can be used for other purposes, such as quality of service (QoS) and deadlock avoidance, by restricting their allocation.

Fig. 2.5 demonstrates the benefits of using VCs while keeping buffer size constant. In this example, the router input and output buffers have the size of four and two flits, respectively. The first scenario shows wormhole flow control where packet A is currently blocked, keeping the physical link between the two routers ideal and blocking packet B.

In the second scenario, both input and output buffers are divided into two VCs. Consequently, packet B can bypass packet A similar to when a car bypasses another slow car using another street lane. As a result, many router micro-architectures adopt this technique for flow control.

#### 2.2.4 Routing Algorithms

The routing algorithms represent the process of calculating and determining the packets path from their sources to their destinations. In router design, the routing unit consists of two functions: routing and selection. The routing function calculates the possible output ports based on packet routing informations such as the destination address and the input port. These options then enter the selection function, which selects one of them either randomly or based on routing information and/or network status.

Routing algorithms can be classified based on several criteria. In this thesis, the most common criteria are briefly listed below along with their pros and cons:

- Based on the number of destinations: routing algorithms could be for unicast (one-to-one ( $1\text{-to-}1$ )), multicast (one-to-many ( $1\text{-to-}M$ )) or broadcast (one-to-all ( $1\text{-to-all}$ )). These different types of traffic exist in networks in different proportions depending on the running application and used protocols. Unicast routing algorithms can be used for all other types if the message is to be duplicated to all destinations. This is known as software multicast, which significantly overloads the network even with small multicast PIR.
- Based on adaptivity: the routing algorithm could be deterministic, partially adaptive or fully adaptive. The first type is when the route from each source to every destination is already determined regardless of network status. As a result, router complexity will be reduced. In contrast, fully adaptive routing algorithms allow the option of path diversity based on network status such as the existence of congestion and faulty links. Although this makes the network more efficient in cases of non-uniform or bursty traffic, this might require very complex router design. On the other hand, partially adaptive routing algorithms try to achieve a compromise between the merits of path flexibility and router complexity by allowing only subset of routes to be selected.
- Based on routing decision maker: this criterion is used to classify the routing algorithms mainly as source routing, centralized routing, and distributed routing. In source routing, the packet path is determined from its source, which makes it mostly deterministic routing. Therefore, routing complexity is reduced

but the packet overhead is not scalable. Centralized routing is very rare due to its lack of efficiency. This is due to the necessity for the centralized unit to distribute routing decisions to all network nodes. In contrast, distributed routing is considered the most favoured type of routing algorithms, where slightly more complex routers collectively determine the packet's path to its destination.

- Based on minimality: some adaptive routing algorithms can be either restricted to minimal paths, called minimal routing algorithms, or are fully free to choose any path, called non-minimal routing algorithm. The former type has lower path diversity and less freedom to adapt to network changing status such as faulty link. However, minimal routing tries to use the least possible network resources and avoids further complexity in the design to avoid livelock and deadlock.

There are other criteria that can be used to categorize routing algorithms, such as whether or not they are oblivious and based on the implementation techniques used [3, 51]. In the next section, turn model techniques are discussed as an example of popular simple routing algorithms.

#### 2.2.4.1 Turn Model

The turn model was firstly proposed by Glass et al. [54] and these techniques are among the most common NoC routing algorithms. This is due to the fact that they are simple, unicast, distributed, minimal, and deadlock-free routing algorithm for two-dimensional (2D) mesh network. Basically, turn-model techniques avoid deadlock by prohibiting some turns of the total possible turns, which are eight. Fig. 2.6 shows a range of turn-model routing algorithms.

The XY turn mode, or dimension-ordered routing (DOR), prohibits four turns as shown in Fig. 2.6a and is therefore considered a deterministic routing. On the other hand, negative-first, west-first, and north-last allow six turns while still offering deadlock-free routing, as shown in Fig. 2.6b, Fig. 2.6c, and Fig. 2.6d, respectively. Therefore, they show more adaptivity than the XY turn model.

Furthermore, Chiu has proposed another turn model that offers higher path diversity, which is odd-even [55]. This routing algorithm has two sets of rules for prohibited and allowed turns, one for odd columns and another for even columns, as seen in Fig. 2.6e and Fig. 2.6f. In addition, inspired by the odd-even model, Dahir et al. have proposed a three-dimensional (3D) odd-even routing that balances the traffic on odd and even XY planes as well as within the planes [56].

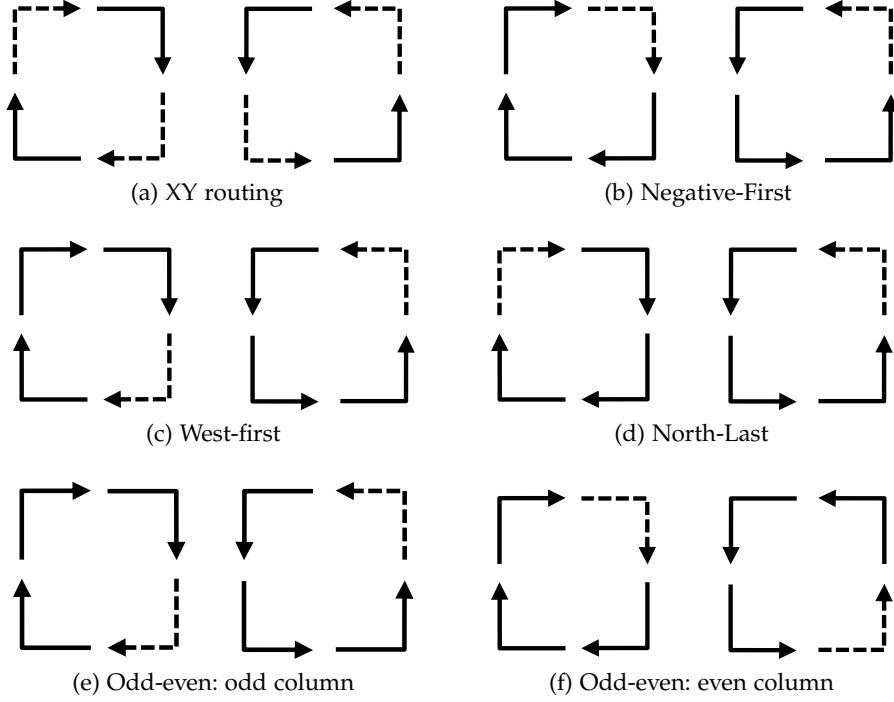


Figure 2.6: illustration of the allowed (solid arrows) and prohibited (dashed arrows) turns in different turn-model routing algorithms in 2D mesh NoCs

### 2.2.5 Arbitration and Allocation

Arbiters and allocators are the building blocks of the control plane in the routers. Arbiters are control units that resolve conflicts over shared resources such as buffers and channels. An arbitration process might be required every cycle, over fixed period of time, or until the requested agent releases the resources. On the other hand, allocators are required to match a set of request agents with a set of resources. There are two main properties of arbiters and allocators that define their effectiveness:

- **Fairness:** this is a property of arbiters, which means that the average number of requests granted over the number of arbitrations for every request agent is the same. Otherwise, a starvation issue might rise. Although this property is preferable, different applications might require different levels of fairness. For instance, some applications require fixed priority or weighted fairness in order to grants some requests more than others.
- **Legal matching:** this is a property of allocators. It means there is no output resource has been granted for more than one input request and there is no input request has been granted more than one resources simultaneously. A maximal match is a favourable feature since it implies maximizing utilization by assigning the

maximum number of legal matches between resources and request agents.

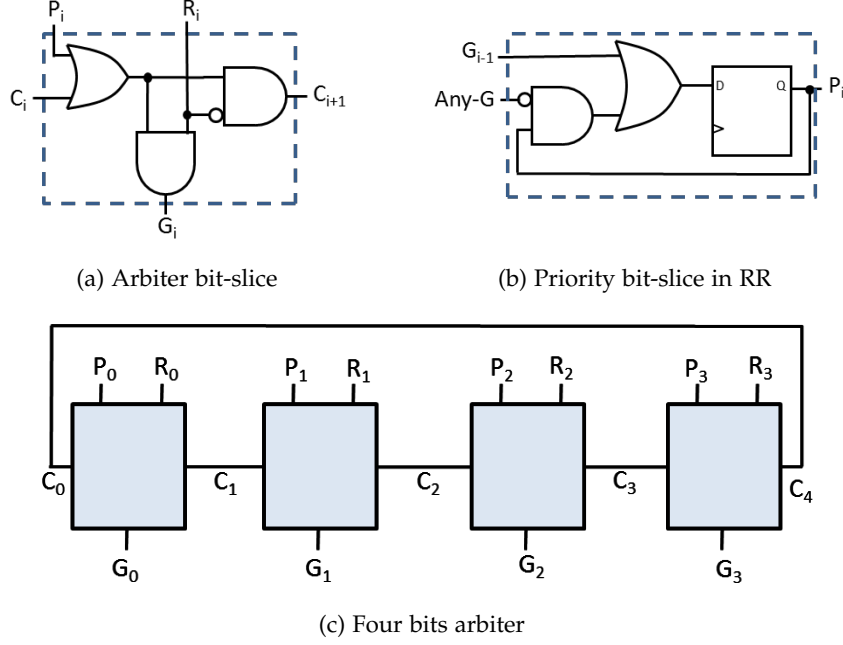


Figure 2.7: Examples of basic arbiters, where R is request, P is priority, C is a carry, G is grant, and  $\text{Any-G} = \bigvee (G_0 \dots G_{n-1})$ .

Since arbiters are considered the most important components of allocators, this section discusses a few basic examples of arbiters. Fig. 2.7c shows a typical arbiter for four requests, in other words a four-bit request vector. The arbiter bit-slice is presented in Fig. 2.7a. This arbiter has priority and request inputs that determine the output grant signal. The arbiter type also depends on the priority feed, which should set only one of the priority vectors at any time. For instance, if the priority is statically set for one of the requests, this is called a fixed priority arbiter, which has no fairness. However, if the arbiter is connected to 4-bit shift register with only one bit set, then it is called a circular or oblivious arbiter. This type of arbiters have weak fairness, since they are unaware of the last request that wins the arbitration. On the other hand, the round robin (RR) type has strong fairness since it enforces the rule, which states that the last winner's request has the lower priority. Fig. 2.7b shows a bit-slice of the priority control circuit of RR. The signal  $\text{Any-G} = \bigvee (G_0 \dots G_{n-1})$  allows priority values to be changed only when a grant signal is set. Moreover, other types of arbiters have been presented in the literature to deal with different arbitration and allocation challenges in NoCs [3, 51, 57].



### 2.2.6 Multi/Many-Core Processors

Since 1971, when microprocessors were invented [58], performance improvements from one generation of processors to the next have been governed by Pollack's law [11]. This states that a performance improvement is equal to the square root of complexity (or area, assuming that the implementation uses the same CMOS technology). In other words, if we double the number of transistors, the resulting processor would have a performance improvements of 40%.

Nonetheless, due to the technology scaling driven by Moore's law, a multi-processor design possibility has emerged which overcomes Pollack's law. This is due to the fact that using multiple processors can offer near-linear performance improvements. For instance, if the number of transistors is kept constant, a double-core processor can offer performance improvements of 70-80% compared to just 40% offered by the double-complexity single large processor [11, 58]. As a result, in the last 15 years, the design of processors has shifted to multi-core designs. Consequently, CMPs and multiprocessor systems-on-chip (MPSoCs) have been developed and adopted by industry [59, 60, 6, 6].

Moreover, due to projected high sub-threshold leakage and the low scaling of supply voltage, the power consumption of the die is becoming impractical [61]. As a result, the clocking frequency, which is the key element of performance speed-up for a single processor, will scale relatively slower than it used to be. The use of smaller cores with low or moderate frequency also solves this problem since the reduction of the size of the core is linearly associated with power reduction while performance is only reduced by the square root of area. Therefore, design trends have moved from multi-core (~10's of cores) to many-core (~100's of cores) since linear performance improvements can then be offered within an affordable die power budget [11].

However, Amdahl's law states that performance improvements are limited by the serial portion of the running code [11]. Consequently, even small percentage of serial code in running application can question the effectiveness of increasing the number of cores beyond certain limit [11, 58]. This limitation is not a considerable challenge, since there are usually many applications running simultaneously with many threads. This almost ensures a good harvesting capability of the computing power of many-core processors.

In addition to the previously stated main benefits of many-core systems, they have extra design advantages such as:

- Each core can be turned on, off, or be idle, thus saving power.
- Using dynamic voltage and frequency scaling (DVFS), individual cores can be tuned to the optimum voltage and frequency, which reduces power consumption. Moreover, the use of heterogeneous many-core systems can be tailored to application requirements and power limitations.

- Many-core systems are inherently reliable since any core could replace a faulty core and/or monitor any failure. Moreover, a load balancing mechanism that distributes heat across the die can improve reliability.
- Many-core platforms speed up the time-to-market since there is no need to design a new complex core. Instead, IP/core reusability and plug-and-play design approaches are possible.

Consequently, many-cores have started to dominate the future of computing. Moreover, the intra-core communication is considered to be the backbone of such systems.

### 2.2.7 Cache Coherence

In all multi/many-core systems, the cores share a main memory and, in most cases, an on-chip last-level-cache (LLC) [62]. However, for each core there is usually a private cache. In this manner, the loading/storing of data from/in the main memory or LLC will be significantly minimized and therefore system performance will improve [63]. In addition, such structures comply with limited power budgets yet utilize the ability to integrate large numbers of transistors, since memory transistors consume less power [11]. Fig. 2.8 presents the structure of such distributed cache systems. Other changes or improvements to this baseline architecture can be adopted, such as multi-level caches, separate cache IPs, or a cache shared by a subgroup of cores. However, this simple tiled architecture satisfies the purpose of this section.

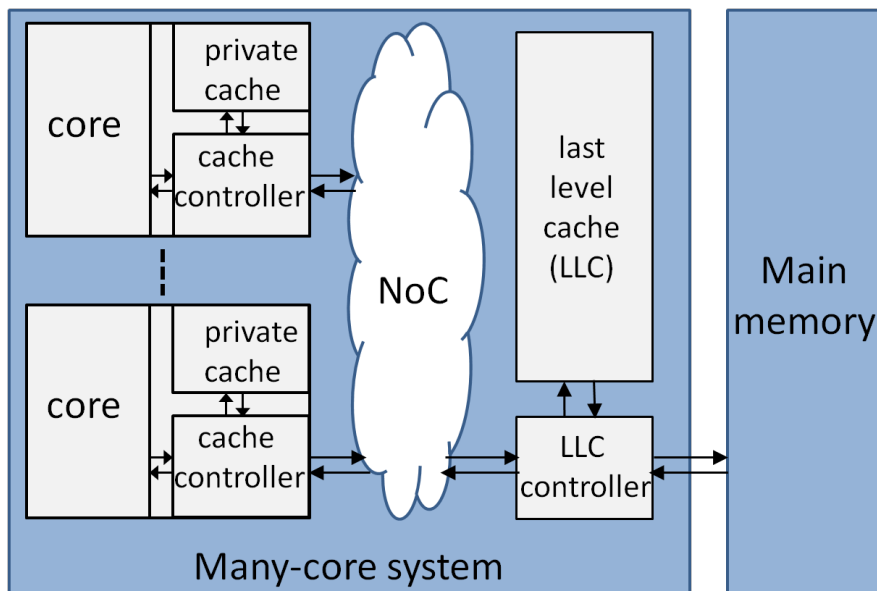


Figure 2.8: Structure of distributed cache for many multi/many-core processors.

This approach of distributed private caches means that there are possibly a multiple copies of the same data block across the chip. Consequently, if a core updates a value of specific data in the private cache, it will be globally inconsistent with other copies of this data and cause incorrect computation. A cache coherence protocol is a technique that ensures the coherency of data across multi-/many-core systems by regulating access to the cache data. This is achieved by enforcing a set of rules (including data status and transactions) that aim to have a single writer yet with possible multiple readers for any data bank at any given time [62].

Cache coherence protocols can be classified into two classes. The first class is snooping, also known as broadcast-based. This class of protocols depends on broadcasting to all other cores request for a data whenever it is needed. Then, the owner sends the data or permission to update its value. As a result, this type of protocols requires a high fan-out feature from the under-layer interconnects. The second class is directory-based, where each cache controller keeps track of the status, owner and/or location of every data block. Therefore, the request is either send as a unicast or multicast to the set of current sharers. This is more complex in terms of functionality and the size of the directory required and therefore it does not scale well with number of cores. On the other hand, broadcast-based protocols are simpler but require excessive communication and are therefore also limited by communication scalability as the number of cores increases. Therefore, currently balancing the pros and cons of coherency techniques is an active research area [63, 62].

#### 2.2.8 Three-Dimensional Integration

Three-dimensional integrated circuits (3D-ICs) are technologies that enable the stacking of 2D dies and linking them together. Many techniques have been proposed, especially in the last decade, to connect die-to-die, die-to-wafer and wafer-to-wafer [64]. However, the most common techniques are either by micro-bonding or through-silicon-vias (TSVs). Recent approaches have suggested wireless vertical communication, such as capacitance and inductive coupling [65].

3D integration is considered to be very attractive approach due to its ability to match circuit scaling density requirements. In addition, it allows the integration of heterogeneous systems with different fabrication technologies such as data memory, photo-electronics, and RF-circuitry. In terms of on-chip interconnects, it is even more appealing for two important reasons. The first is a significant reduction in interconnect length across the chip. This reduction is proportional to  $\sqrt{N_{\text{dies}}}$  [64], where  $N_{\text{dies}}$  is the number of stacked dies, assuming that the total area remains constant. The second reason is a reduction in the hop-count and the increased path diversity of packets through

the NoC routers [66]. However, even though the vertical links are much shorter than horizontal links, their low density limits die-to-die bandwidth [64, 67].

Moreover, despite all the advantages mentioned above, 3D integration faces various technical challenges such as process control requirements, wafer thinning, packaging, low through-silicon-via (TSV) capacitance, high fabrication costs, and design challenges [64, 9, 7, 10]. However, the thermal impact of stacked dies, which are very dense and not directly attached to a heat sink, might be the biggest challenge facing this technology [64, 67]. This thermal impact will decrease performance of the integrated circuits and their life-time. As a result, this is at the moment considered to be a not fully matured technology and an active research area. However, great progress is being achieved in these areas and a number of solutions can be offered for each problem [7, 64]. However, 3D-NoCs are not considered further in this thesis since they are beyond the scope of this study.

## 2.3 LITERATURE REVIEW

### 2.3.1 *Current and Emerging Interconnects*

This section discusses a set of promising emerging types of interconnect. These interconnects have been proposed in the literature as replacement or supplement interconnects to regular wire, which projections show will soon face serious issues in terms of global and multicast communication.

#### 2.3.1.1 *Wire Issues*

The on-chip interconnect trend for decades has been to rely only on the regular metal wire interconnect, which transmits the signal by charging/discharging the whole wire. These wires, also known as resistive and capacitance (RC)-lines, provide a cheap and easy to implement communication medium. Although the interconnect fabric has changed from bus to NoC [51, 3], the under-layer media is still the same. Wires have been meeting all of the performance, power consumption, and economical implementation requirements for intra-chip communication for many generations of the technology. However, with the continuous scaling of complementary metal-oxide-semiconductor (CMOS) technology, the projections for wired global communications do not seem promising.

Even though global wiring length might remain the same or increase slightly, wire thickness and spacing have been continuously decreasing as technology has scaled down. This increases wire resistance and capacitance [7, 10, 68]. Subsequently, wire delay increases because it is inversely proportional to wire resistance and capacitance [1]. Fig. 2.9 shows the increasing gap between gate delay and wire delay [7].

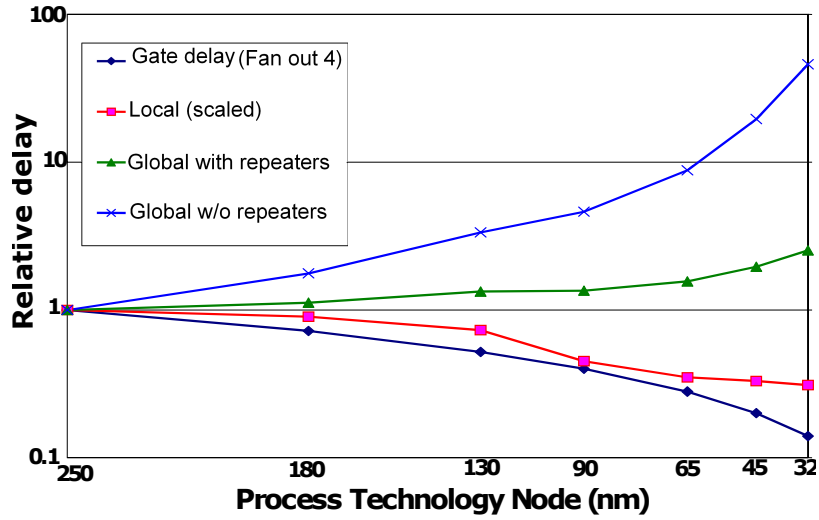


Figure 2.9: Projected delay issues of global regular wire compared to gate delay with technology scaling [7].

Moreover, Ho et al. predicted that global and semi-global wiring delay for delivering 50% of the signal (threshold of transferring from logic 0 to logic 1) might exponentially increase [8]. As a rule of thumb this delay equal to  $0.4d^2RC$ , where  $d$  is the distance, and  $R$  and  $C$  are the resistance and capacitance of wire per meter, respectively. Local wiring does not have this problem because, unlike global wiring; its length decreases with technology scaling as shown in Fig. 2.9.

This latency will decrease the single wire bandwidth and overall interconnect throughput. The attempt to keep wire dimensions (thickness and spacing) constant regardless of technology scaling is known as fat wires. This approach has a serious drawback, which is to reduce the ratio of bits per area and thereby aggregated bandwidth could be severely reduced comparing to the delay resulting from decreasing the wire geometry [8]. However, the industrial sector now uses mixed wire geometry in different layers in the IC based on wire length and functionality in order to mitigate the delay problems [10, 68]. Other solutions, such as introducing a new conductor and dielectric material with better physical characteristics [10] or using repeaters [69, 7], could postpone the problem for a few years but will be unable to meet future demands. For instance, some studies show that introducing repeaters for global and semi-global wires mitigates the delay problem by making the delay rise linearly with technology scaling [69, 7].

### 2.3.1.2 Optical Interconnects

Optical-based interconnects or optical networks-on-chip (ONoCs) offer many significant features which would overcome the drawbacks of wires in global communications, such as high bandwidth per channel, low electromagnetic interference, the ability to cover longer distances, and speed-of-light signal propagation [27, 28, 29, 30, 7, 70]. These fea-

Table 2.1: Summary of reported key features for implementation of integrated optical interconnects.

study	technology node	Gb/s	pJ/bit
Meade et al.[71]	180nm	5	2.8
Dong et al.[29]	-	20	-
Cunningham et al.[72]	40nm and 130nm	10	0.344 without optical source and ring tuning power
Zheng et al.[73]	40nm	10	0.53 without optical source and ring tuning power
Faralli et al.[74]	220nm	10	-

tures have led researchers to investigate this type of interconnect for on-die communication given that it has previously been limited to relatively long range, such as on-board communication or longer. For instance, to utilize speed-of-light communication to eliminate clock skew, the optical interconnect has been proposed for clock signal distribution [7, 70]. Moreover, its projected high aggregated bandwidth (up to 1Tb/s) could satisfy the future large intra-chip data communications of many-cores [28]. Recent work has achieved up to 20Gb/s per physical channel [29]. In terms of maturity, significant advances have been demonstrated in the last few years in silicon photonics, as shown in Table 2.1.

Despite all of the previously mentioned merits, optical interconnects face significant challenges, mainly in terms of complexity, thermal regulation, and power budget requirements [28, 75, 76]. In terms of power consumption, there is debate over whether optical interconnects will reduce or increase overall power consumption. Many optimistic researchers [27, 77] argue that the absence of resistance loss and the assumption that quantum sourcing and detecting can be used in the future could offer better bit/J than regular metal wires. In contrast, pessimistic researchers question the potential power savings unless these interconnects are used for relatively long communication distances, since currently proposed optical devices are so power-hungry [30, 10]. Moreover, researchers have yet to tackle the extra power requirements for scalable multicast, since current devices decay the signal significantly.

Optical interconnects have another major challenge, which is their complexity. For instance, they need expensive area-hungry and sometimes non-CMOS devices to transfer signals from electrical to optical form and for routing optical signals [27, 28]. The main devices used are laser sources, photo detectors, modulators/filters, waveguides and

laser-waveguide couplers in the case of off-die laser sources. Also, depending on the interconnect architecture, other optical devices might be needed such as nanoscale mirrors, micro-lenses, photonic switching elements and splitters/combiners. Some of these devices, such as laser sources, might need to be placed off-chip [30, 10]. This creates issues with manufacturing complexity (such as packaging and pin number requirements) and high coupling losses that might dominate the power consumption budget [10]. However, advances in more-than-Moore options represented by silicon photonics devices have almost eliminated the CMOS compatibility challenge. This is achieved by developing techniques to integrate almost all optical devices on silicon chips such as the silicon-germanium photodetector and polysilicon optical modulators [71, 78, 10]. However, some of these techniques are not yet mature and face their own challenges. For instance, the columnar polysilicon optical waveguide has attenuation loss of 40dB/cm [71]. In general, optical interconnects are still a relatively costly alternative despite all the significant advances in the last decade in silicon photonics.

On the other hand, optical waveguide routing is constrained so that no hard turns are allowed in order to avoid major signal degradation. Other major challenges include careful thermal tuning management, which is required in microring-based wavelength filters otherwise thermal variation might lead to link failures [29, 75, 79]. This could be difficult in dense VLSI applications such as future many-cores with variable switching rates and/or in harsh environments. Moreover, optical devices with critical dimensions are found to be sensitive to integration process variation, which is a natural result of CMOS fabrication [79]. Both thermal and process variation cause the passband of the optical transmitter and receiver to mismatch and this leads to signal loss and crosstalk [79]. As a result, some techniques have been introduced to mitigate thermal and process variation, but their complexity or power overheads increase the optical network-on-chip (ONoC) existing cost and power budget challenges [79]. Therefore, these challenges are likely to prevent the optical interconnect from being preferred in the near future.

### 2.3.1.3 *Wireless Interconnects (WiNoC)*

RF-based interconnects such as wireless interconnects or wireless NoCs (WiNoCs) appear to be a cost-effective alternative compared to optical interconnects [22, 23, 9]. This is due to the fact that radio frequency (RF) circuitry is compatible with CMOS technology and therefore it is less area and power-hungry. Many studies have proposed wireless network-on-chip (WiNoC) solutions as either supplementary [24, 25, 26, 80] or possibly replacement [81] interconnects for regular wire-based NoCs. This type of interconnect basically transfers the electrical signal into an electromagnetic (EM) signal via the use of an integrated transceiver and antenna. This EM signal would propagate in

Table 2.2: Examples of demonstrated integrated wireless communication systems along with their key features for a single link.

study	technology node	modulation	Gb/s	pJ/bit
Chen et al.[84]	180nm	ASK	6	17
Wang et al.[85]	90nm	FSK	1	-
Yu et al.[86]	65nm	OOK	16	-
Kawasaki et al.[87]	40nm	ASK	11	6.4
Okada et al.[88]	65nm and 40nm	16QAM and QPSK	3-6	11.8
Kawai et al.[89]	65nm	16QAM	7	-

one-hop via free space to the surrounding nodes in the coverage area at nearly the speed of light. In terms of physical channel bandwidth, predictions show an increase in transistor switching speed as CMOS technology scales down. This would enable the use of higher carrier frequencies [82, 7, 83]. As a result, a wide spectrum of frequencies up to the terahertz (THz) is possible, which is necessary to allow multi-channel realization at this shared medium [22]. Moreover, these high frequencies would require an integrated antenna of smaller size. Table 2.2 reviews examples from the literature of implemented integrated wireless communication systems. This wide range of studies shows the level of technology maturity of this type of interconnect.

WiNoC technology is considered to be one of the most mature emerging interconnect types since many implementations of WiNoC components such as integrated antennae and transceivers have been presented in the literature [90, 81, 86, 91]. However, so far, there are some challenges facing WiNoC. For instance, researchers are finding it difficult to design an antenna with wide frequency bandwidth, low power dissipation, larger coverage area and small area overhead [22]. Firstly, the WiNoC channel bandwidth is limited by the antenna operational frequency ( $F_c$ , the central resonance frequency) and the 3 dB bandwidth (B). For example, the 0.38 mm zigzag antenna has a transmission gain ( $S_{21}$ ) showing B around 15 GHz [92]. Antenna percentage bandwidth ( $B_r$ ) is inversely proportional to operational frequencies, as shown by the following equation:

$$B_r = \frac{F_1 - F_2}{F_c} \times 100\% \quad (2.1)$$

where  $F_1$  and  $F_2$  are the starting and ending frequencies respectively of the 3 dB bandwidth. For example, the zigzag antenna mentioned earlier has  $B_r = 27\%$  [92, 22]. Thus, the WiNoC link might require a cluster of antennae with different central frequency and design charac-



teristics in order to collectively provide the required frequency range. Other solutions include the use of antennae with high operational frequencies, such as in the THz range, where they would consume less area and have wider frequency bandwidth [93]. However, these solutions waste a large part of the available frequency spectrum, which is limited and governed by the CMOS technology cut-off frequency. The second solution is a time multiplexing approach [94], which obviously decays the throughput of the channel. In terms of area overheads, integrated antennae are considered to be area-hungry passive components [22, 94]. However, antenna dimension is inversely proportional to operational frequency. Therefore, with the scaling down of technology and the realization of THz, the area overhead could be effectively reduced [22, 94, 90]. Other solutions for antennae include the use of carbon nanotube [95] or planer graphene [23]. These techniques could improve power and area budgets and might allow to some extent a configurable resonance frequency. However, the implementation challenges of these technologies have yet to be addressed.

Other challenges facing WiNoC are related to channel reliability. Due to the nearby circuitry, noise could be injected into the transceivers or the antennae [90]. However, previous studies show that effective isotropic radiated power (EIRP) has almost negligible effects on adjacent circuits such as DRAMs [96] and analog-to-digital converters [97]. Moreover, many studies have addressed the alleviation of channel interference and error rates by adjusting transmitter power, in other words, adjusting the signal-to-noise-ratio (SNR) [98, 25]. Other reliability issues result from the antenna being influenced by the chip packaging [90]. Therefore, these issues need to be carefully considered in transceiver and antenna design in order to become an efficient design option.

#### 2.3.1.4 RF-interconnects with Transmission Lines (RF-I)

The other alternative to electromagnetic free space signal propagation is waveguided propagation via transmission-lines (TLs), which is known as the RF-interconnect with transmission-line (RF-I) [82, 19, 20, 18, 21]. These types of interconnects are similar to the WiNoC in terms of CMOS compatibility, signal velocity close to the speed of light, low global communication energy and high throughput compared to regular wires. As a result, many studies have proposed RF-I as a supplementary interconnect for the metal wire [82, 18]. Moreover, some researchers have even discussed the possibility of replacing metal wire with RF-I [19]. These studies utilize the RF-I either as a special-purpose interconnect [19, 99] or as general purpose express links [82, 18]. In terms of RF-I maturity, demonstrations of on-chip RF-I implementations have been presented in many studies [98, 21, 100, 101]. Table 2.3 presents the key features of some recent on-chip implementations

Table 2.3: Examples of reported implementations of integrated transmission lines along with their key features for a single link.

Study	Technology node	Gb/s	pJ/bit	TL
Chang et al.[98]	180 nm	4-20 (predicted)	-	CPW
Chang et al.[101]	90 nm	5	20	CPW
Hsu et al.[21]	90 nm	-	-	modified CPW
Ito et al.[100]	90nm	8	0.3-0.9	CPS

of RF-I in the literature. Moreover, high-end chips that utilize global transmission lines for clock distribution already exist [99].

RF-interconnects with transmission-lines (RF-Is) require an integrated transceiver, similar to WiNoC, to transform the electrical signal into an RF signal. However, instead of an antenna, the RF-I uses the on-chip transmission lines as waveguides to propagate the signal. Consequently, the RF-I has less power dissipation and less power consumption is required. There are three main types of on-chip TLs [19, 102], which are the microstrip line (MSL), the coplanar waveguide (CPW), and the differential line or coplanar strip (CPS), see Fig. 2.10. The MSL is known for its simplicity compared to the CPS and CPW, while the latter two show better robustness against crosstalk, especially in mm-waves [19]. Moreover, the CPS is known for its higher interconnect density compared to the CPW [19].

RF-I inherit the same limitation as WiNoC in terms of interference and the cut-off frequency of the CMOS technology. However, in terms of limited frequency, designers have the option to have more than one shared media by adding more TLs. This would increase the aggregated data bandwidth [19, 82]. Moreover, unlike with the WiNoC, the frequency spectrum of RF-Is is not limited by the resonance frequency of the antenna and  $B_r$ .

In terms of interference and channel reliability, the main challenge facing the RF-I is crosstalk among TLs and between TLs and the surrounding circuitry. This is especially true at high frequencies or in long TLs [22, 103], and is due to increasing resistivity because of the skin-effect (reduces the cross section) that increases with the increase of operational frequency [20]. As a result, many studies have proposed various techniques to improve crosstalk robustness, such as designing the TLs with low impedance ( $Z_0$ ) [100]. However, this will increase the power dissipation [100]. Other studies mitigate this issue by proposing power and ground shielding lines between the TLs [100], increase spacing between TLs [104], adopt or develop special TLs types such as

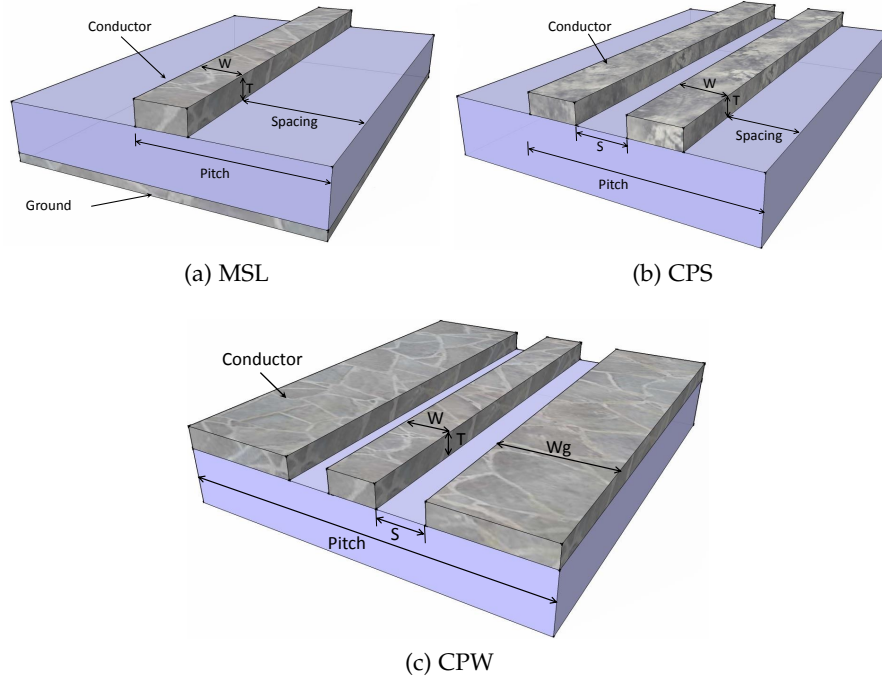


Figure 2.10: Structure of the main three types of the transmissions lines: (a) microstrip line (MSL), (b) differential line or coplanar strips (CPS), and (c) coplanar waveguide (CPW).

CPW and CPS. However, these suggested solutions increase complexity and decrease the aggregated bandwidth per area.

The second main challenge concerns the area overhead and interconnect density. These TLs are fabricated using the upper-layer of CMOS metal wires because of the thickness required. This is due to the fact that these high dimension wires have low resistance. However, they have large capacitance and therefore require a wider inter-metal dielectric to mitigate parasitic capacitance [20]. Moreover, some studies propose the insertion of a metal pattern underneath the transmission lines in a multi-layer design to reduce parasitic effects and crosstalk [21]. These costly wires might need to span the whole chip in a worm or cycle layout, as mentioned earlier. Thus, significant performance improvements of TLs would be required to justify this extra cost.

The third main challenge is the limitation of drop points, which raises the question of the level of RF-I scalability in many-core processors with hundreds or thousands of cores [22, 103]. These drop points are necessary to fully utilize these costly fat wires by having a multichannel frequency instead of many segments of TLs [19, 82]. In addition, these multi-drop points are needed to providing the fanout feature, as mentioned earlier. Therefore, many researchers have tried to mitigate multi-drop scalability and transmission-line (TL) discontinuity [19, 82, 100].

### 2.3.1.5 Surface Wave (SW)

The Zenneck surface-wave (SW) is an inhomogeneous EM wave supported by a metal-dielectric surface. The designed surface is a waveguide that traps the EM signal in a 2D medium instead of three-dimensional free space. As a result, the E-field decay rate in the SWI from the source horizontally along the boundary is around  $(1/\sqrt{d})$ , as shown in Fig. 2.11, where  $d$  is the distance from the source [105]. This feature allows the Zenneck surface-wave interconnect (SWI) to offer relatively linear J/bit over this short distance compared to the high scaling of regular global buffered wire interconnects.

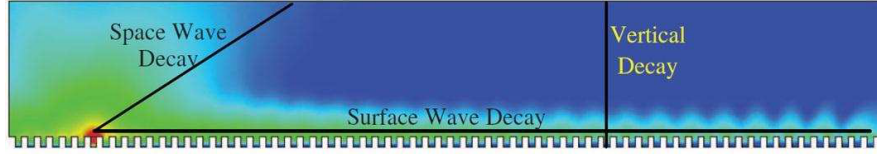


Figure 2.11: Zenneck surface wave propagation decay which is significantly better than free space propagation [105].

The surface wave communication can be realized in various ways. The surface waveguide should be engineered by altering its dimensions, and the materials used for the conductor and/or dielectric. These characteristics are chosen so that the characteristic impedance ( $Z_0$ ) will be around  $(10 + j300) \Omega$ . For instance, the designed surface medium could consist of a dielectric layer placed over a corrugated conductor layer [106, 105, 107, 31]. According to Hendry et al., the surface characteristic impedance of the corrugated surface can be calculated using the following equation:

$$Z_s = jZ_w \frac{d_2}{d_1} \tan\left(\frac{2\pi d_3}{\lambda}\right) \quad (2.2)$$

where  $d_1, d_2$  and  $d_3$  are defined in Fig. 2.12,  $Z_w$  is the characteristic impedance of the dielectric material in the grooves, and  $\lambda$  is the wavelength. The dimensions of the corrugated surface are proportional to the wavelength, with at least three or four repetitions in one wavelength and a depth less than quarter of the wavelengths [105]. Moreover, a dielectric-coated metal flat surface has been developed successfully using dielectric material, which has the required impedance, overlaid with a conductor sheet [32].

Moreover, in order to have a Zenneck surface wave, a positive reactance ( $X_s$ ) is required [106, 32], where  $X_s$  is derived as follows:

$$X_s = 2\pi f \mu_0 \left[ \frac{\epsilon_r - 1}{\epsilon_r} l + \sqrt{\frac{1}{4\pi f \mu_0 \sigma}} \right] \quad (2.3)$$

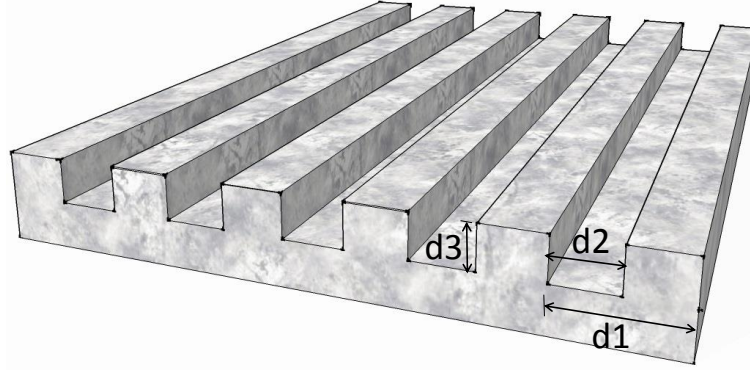


Figure 2.12: The corrugated surface.

This equation shows that reactance depends on the operational frequency ( $f$ ), dielectric constant or permittivity ( $\epsilon_r$ ), thickness of the dielectric layer ( $l$ ), permeability of the free space ( $\mu_0$ ), and the conductivity of the metal layer ( $\sigma$ ). More details of surface wave propagation theory can be found elsewhere [106, 105, 107].

On the other hand, maximum transmission into the SW occurs when the incoming wave is incident at or close to the Brewster angle, where reflections are minimized. Therefore, the integration of a transducer linked to the transceiver is needed to launch the waved signal into the surface [106, 108]. This can be as simple as, for omni-directional transmission, a coaxial to parallel plate (or flange) waveguide [105]. Also, it could be a dipole or monopole for omni-directional communication, with a parallel plate waveguide [109].

Recently, experiments have demonstrated the transfer of data using two coaxial waveguide transducers, one for the receiver and the other for the transmitter, with a corrugated aluminium sheet as surface wave guides between them [106]. Measurements were conducted in the frequency range 10 GHz-50 GHz for distances up to 1.26m and simulations were carried out up to 100 GHz using HFSS and CST simulators [110, 111]. Unlike the WiNoC, there is no limitation of frequency range as long as the appropriate transceiver and dielectric material are used. For instance, Fig. 2.13 shows simulation results for a flat surface and 50 mm distance. Obviously, these results demonstrate that if careful consideration is given to the choice of dielectric material for the waveguide surface, the surface can carry a wide range of frequencies with relatively very low signal dissipation.

In terms of reliability, due to the lossy free space propagation of WiNoC, the bit error rate (BER) of SWI ( $10^{-14}$ ) is much better than the BER of WiNoC ( $10^{-7}$ ) for the same signal power [32]. Therefore, since the BER SWI is comparable to wire BER, SWI can achieve performance improvements up to 30% over WiNoC [32, 112].

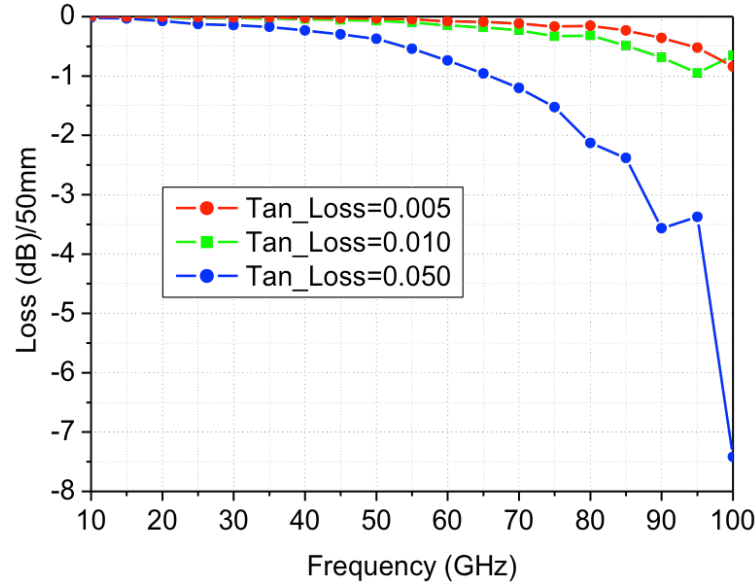
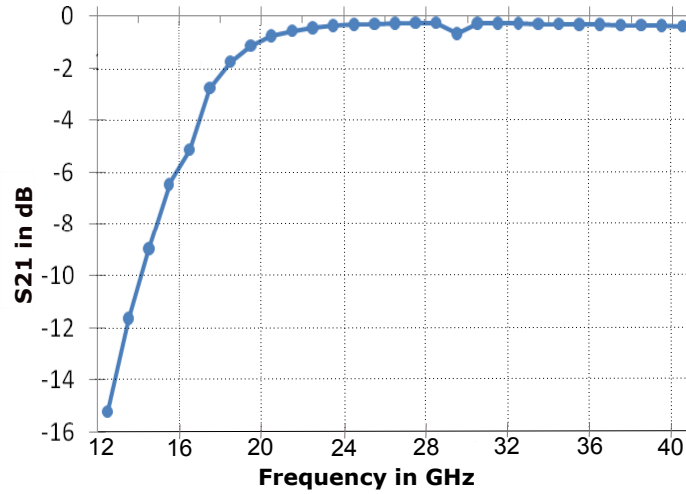


Figure 2.13: Flat surface wave signal power loss proportional to frequency in the case of a dielectric material with relatively high loss tangent (Tan-Loss). These results obtained in collaboration with K.Tong in University Collage of London

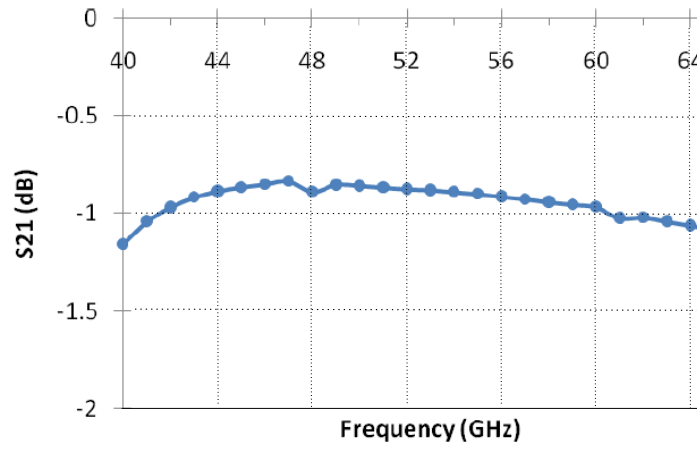
#### 2.3.1.6 Comparative Summary

Table 2.4 presents a summary comparison of the key features that will be crucial in future interconnect architectures. Power consumption is the main limitation for future interconnects, especially after projections which show that interconnect fabrics might consume a non-trivial percentage of the entire chip's power consumption [113]. As shown in Table 2.4, RF-based interconnects that use waveguides have relatively low power consumption since they neither require power-hungry devices nor involve high power dissipation.

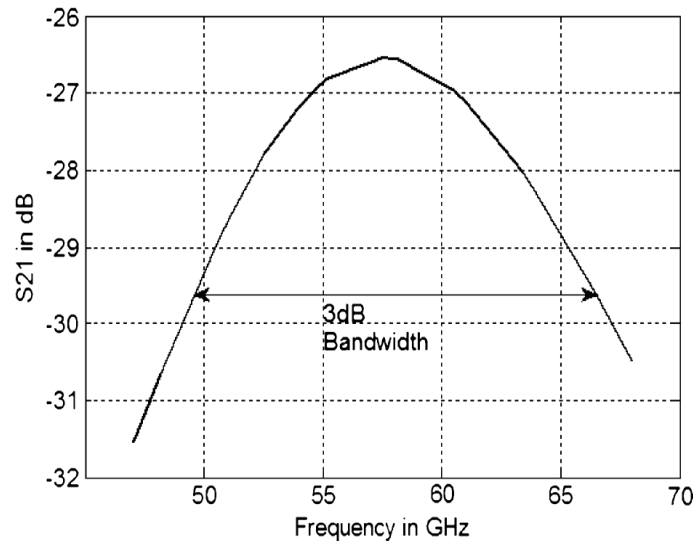
In terms of signal decay and reliability, the signal integrity of optical interconnects is superior to other interconnects. The second best type in terms of reliability is the SWI. This is due to the fact that, unlike the RF-I, the designed surface waveguide is almost immune to interference from nearby circuitry. On the other hand, the WiNoC and SWI show remarkable natural fanout features compared to other emerging interconnect types. This feature is crucial for scalable multicast architectures in future many-cores processors, especially since 1-to-M and 1-to-all traffic PIR, size, and the likelihood of creating hotspots could increase with the increase of number of cores, as will be discussed further in chapter 5.



(a) Surface-wave (low frequencies)



(b) Surface-wave (high frequencies)



(c) Zigzag antenna

Figure 2.14: Comparison of forward transmission gain ( $S_{21}$ ) between Wireless [22] and SW interconnects [105].

Table 2.4: Summary comparison of key features in current and emerging on-chip interconnects.

Features	Metal wire [7, 8]	Transmission lines (RF-I) [82, 19]	Wireless connect [22, 24, 7]	Optical interconnect [28, 27, 77, 7]	Surface wave interconnect (SWI)
<b>Power</b>	Dynamic power that is proportional to the wire capacitance and voltage.	Power consumption is relatively tolerable.	High free space power dissipation.	High power consumption.	Power consumption is relatively tolerable.
<b>Signal Decay</b>	Limited by latency, which increases exponentially without repeaters.	Low signal decay and dissipation.	High decay, inversely proportional to distance.	Very low signal decay and dissipation	Low signal decay and dissipation inversely proportional to square root of distance.
<b>Reliability</b>	Possible crosstalk.	Crosstalk exist (capacitor and inductor coupling).	Noise coupling to the antenna and possibility of multi-path interference.	High signal integrity.	Less subject to noise coupling.
<b>Fan-out</b>	Needs extra power for multi-drop bus (stubs) and lowers propagation velocity.	Stubs impedance discontinuity, which will lead to signal reflection.	Limited by transmission signal propagation coverage area only	Requires optical splitters and combiners that decay the optical signal (3dB per splitter).	Limited by transmission signal propagation coverage area, which is wider than WiNoC.

... to be continued



Table 2.4: (continued) Summary comparison of key features in current and emerging on-chip interconnects.

Features	Metal wire [7, 8]	Transmission lines (RF-D) [82, 19]	Wireless interconnect (WiNoC) [22, 24, 7]	Optical interconnect [28, 27, 77, 7]	Surface wave interconnect (SWI)
<b>Bandwidth</b>	Limited by interconnect delay; thus, bit rate is dependent on distance.	Limited process technology transistor cutoff frequency, which is currently 100 to 200 Gbps.	Limited process technology transistor cutoff frequency, which is currently 100 to 200 Gbps	Very large bandwidth with multi-wavelength capability up to 500 Gbs.	Limited process technology transistor cutoff frequency, which is currently 100 to 200 Gbps.
<b>Complexity</b>	Low and medium complexity for local and global communication, respectively. Needs repeaters for cross-chip communication that consume transistors and VIA while also restricting floor planning. However, still the cheapest and simplest interconnect.	Medium complexity required: (1) integrated transceiver, (2) wide thick wires and spacing (12-45 $\mu$ m), (3) may require shielding wires and metal planes to overcome coupling, (4) matching circuits in case of forking path	Medium complexity. Requires: (1) integrated transceiver, (2) integrated antenna or cluster of antennae based on the required bandwidth and operational frequency.	High complexity and some devices are not CMOS compatible. Requires: (1) laser source, (2) photo detectors, (3) modulators and filters, (4) waveguide, (5) laser waveguide couplers in cases of off-die laser source, (6) nanoscale mirrors, (7) splitters/combiners.	Medium complexity. Requires: (1) integrated transceiver (2) integrated designed surface (3) integrated transducer.

Given the projected scaling in number of CMPs cores and predicted volumes of their communication, interconnect bandwidth is considered to be one of the main requirements of future many-core processors. All RF-based interconnects are limited by the cut-off frequency of CMOS technology. However, the cut-off frequency will continue scaling up as technology scales down. On the other hand, as mentioned earlier, the operational frequency of antenna and relative bandwidth further limit the bandwidth of the WiNoC data channels. For instance, the 0.38 mm Zigzag antenna has a value of  $B$  around 15 GHz [22], as shown in the reported transmission gain ( $S_{21}$ ) in Fig. 2.14c. In contrast, Fig. 2.14a and Fig. 2.14b show the SWI transmission gain ( $S_{21}$ ) with a much wider frequency spectrum [106, 105]. On the other hand, optical interconnects surpass other emerging interconnects in terms of bandwidth. However, due to CMOS compatibility and the costly implementation process for optical interconnect devices, ONoC is so far an expensive solution as shown in Table 2.4. In contrast, other emerging RF-based interconnects such as the WiNoC, RF-I, and SWI are CMOS-compatible and therefore are relatively cheap devices.

### 2.3.2 Existing NoC Architectures

This section presents examples of selected cutting-edge many-cores systems utilizing NoCs that have been developed in academic and/or industrial contexts.

#### 2.3.2.1 SpiNNaker

The SpiNNaker is a cooperative project between Manchester and Southampton Universities and many other entities [114]. This project aims to build a massive parallel processing architecture with up to one million processing units intended to emulate the human brain. This is due to the fact that spiking-neural-network (SNN) has high parallelism and can harvest the computing power of such a system more efficiently than general purpose applications. However, it has recently been proposed for many other applications such as in robotics and computer science [114].

The building blocks of this architecture is the SpiNNaker multiprocessor chip that consists of 18 ARM968 cores. Two of these cores are used for monitoring and redundancy for fault tolerance while the rest run the application software. These cores are connected to a 128MB SDRAM on top of it using gold wire bonding [115, 116]. Each chip is connected to six neighbour chips, creating a 2D triangular mesh topology. Moreover, these chips are connected to each other using circuit boards; the latest board incorporates 48 chip nodes [115, 114]. The project aims to connect around 1200 boards to achieve its aim.

In terms of interconnects, the SpiNNaker chip has two NoC routers. The first is for both on-chip and off-chip inter-core communication,

also known as the communication NoC. The second connects the cores to the peripherals and memory, and is also known as the system NoC. The router supports multicast communication and is able to replicate the packet to more than one output (tree-based multicast). The routing algorithm is deterministic (application-based) and is implemented using a lookup table (with 1024 words) [117]. This on-chip router represents a crossbar on-chip topology enabling 2-hop communication for intra-chip communication. This design choice, along with the implementation of one turn per direction in the router, was made to reduce the NoC area. However, the NoC area is still non-trivial ( $\sim 10\%$ ). Moreover, this system is not ideal for applications that require communication reliability, since it drops packets after a waiting period without resending them. However, this is an ongoing project and is considered to be the tip of the spear for NoC-based many-core systems.

### 2.3.2.2 Intel Many Integrated Core (MIC) Architectures

In terms of processing units, industry has all but abandoned the single core for multi/many-core processors in attempt to continue to follow Moore's law. The Intel many-integrated-core (MIC) architecture project is an obvious example of this trend [118, 119]. Over the last decade, this projects depicts has yield two prototypes and one product.

The 80-core TeraFlops with  $10 \times 8$  2D mesh NoC was first presented in 2008 [6]. The chip was fabricated using the 65nm technology and it consumes around 100W. The NoC consumes 28% of this power and represents 17% of the die area. The routers specifications include a 16-flit buffer per port and it has five ports connected to one processing element (PE) and four other routers using 32bit links. In addition, it offers 4-cycle per hop latency and operates at 4GHz.

The second prototype, which is the single-chip cloud computing (SCC), was presented in 2011 [5]. This state-of-the-art architecture increases the complexity of the tile to include two cores of Pentium class IA-32 and two L2 caches. The chip has 24 tiles connected via  $6 \times 4$  2D mesh NoC, and thus the total number of cores is 48. Fabricated using 45nm technology, this chip was designed to offer DVFS. In terms of interconnects, the SCC NoC offers 2.8x power/performance improvements over the TeraFlops NoC. This NoC runs at lower operational frequency (2GHz), and consumes 10% and 5% of the chip power and area, respectively. This is remarkable since the wire link width is increased to be 16 bytes and the router ports have eight VCs per port, where each VC can hold one packet (3 flits). These VCs are divided into two message classes, which are request and response, and one of each class is reserved for breaking deadlock cycles. The router uses XY routing and virtual-cut-through switching.

Although these two prototypes were very promising, the first genuine many-cores commercial product offered by Intel was the Xeon Phi coprocessor in 2013 [118]. This chip is manufactured using 22nm

technology and includes up to 61 cores and consumes up to 300W. The cores, memory and other peripherals are connected using a bidirectional ring NoC. Unlike the mesh, this topology can offer uniform communication bandwidth for all cores, since the cores at the edges of a mesh have fewer connections than the rest. Moreover, this NoC topology is more power efficient and costs less for a low number of cores than the mesh. However, the researchers have acknowledged the drawbacks of this design decision [118], which include the following:

- Low scalability, since latency is proportional to the number of cores ( $N/4$  hop latency).
- Not robust to failure and one damaged link can cause the whole chip to be faulty.
- Easily congested network.

The links, however, are four times wider than the SCC (64byte) with a separately controlled bidirectional links. On the other hand, the routing algorithm is either deterministic or the packet continue to bounce until it reaches its destination. Obviously, the interconnect architecture has many known limitations in terms of scalability but it seems to be cost-effective for the current number of cores.

### 2.3.3 NoC Simulators and Models

Due to the need to evaluate and explore the potential of different design parameters and configurations of NoCs and the high cost of NoCs platforms prototyping, many studies have been dedicated to develop NoC simulators, models and design tools. This is especially true for system-level evaluation tools and simulators, since these are essential to match future application requirements to the under-layer communication fabric and to explore the potential and limitations of proposed designs. These simulators either evaluate the performance, power consumption, and/or area of the design [120, 121, 122, 123, 124, 125, 126]. Table 2.5 presents a comparison of some of these simulators along with those developed in this thesis. Obviously, some of these simulators have been further developed to offer more features. For instance, Noxim has been developed to simulate the WiNoC [127, 128]. Moreover, other researchers have adjusted it to simulate 3D mesh/3D routing algorithms [67, 52] and the torus/twisted-torus [129]. In this thesis, the Noxim simulator is adjusted also to evaluate the proposed architecture, as will be discussed in the following chapters.

Table 2.5: Comparison of various NoC models and simulators.

Simulator	Topologies	Routing	Switching	traffic	Interconnects	Evaluated
Booksim [120]	2D mesh, torus, flattened, butterfly, fat-tree, quad-tree...etc.	deterministic, adaptive, minimal, non-minimal...etc.	Wormhole with VCs	Uniform random, bit complement, bit reversal, shuffle, transpose, tornado, neighbour and random permutation	Wire	Performance
Noxim [121, 128]	2D mesh	deterministic, turn models, fully-adaptive	Wormhole	Bit reversal, uniform (random), transpose, shuffle, hotspot, butterfly and table-based	wire, WiNoC	Performance, Power
Orion [122, 123]	Unrestricted	Not applicable	Wormhole with VCs	Not applicable	wire	Router power and area
Netmaker [124]	Unrestricted	Deterministic XY	Wormhole with VCs	Uniform random	wire	Performance
Nirgam [125]	2D mesh and 2D torus	XY, odd-even, and source routing	Wormhole with VCs	Customizable constant bit rate, trace-based and bursty traffic	wire	Performance

... to be continued

Table 2.5 (continued): Comparison of various NoC models and simulators.

Simulator	Topologies	Routing	Switching	traffic	Interconnects	Evaluated
PhoenixSim [126]	crossbar (photonic)/ mesh and torus (wire)	source routing and wavelength-selective routing	circuit-switching (photonic), wormhole with VCs (wire)	random, hot-spot, nearest-neighbour, and tornado	Optical interconnect, wire	Performance and power
Chapter 3, 4	2D mesh, W-SWI	deterministic, turn model and fully-adaptive	Wormhole	Uniform random, bit reversal, shuffle, transpose, hotspot, butterfly and table-based	wire, SWI	performance and power
Chapter 5, 6	2D mesh, W-SWI	deterministic, turn model and fully-adaptive, virtual-circuit-tree (VCT), wire and surface-wave interconnects (W-SWI) multicast routing	Wormhole with VCs, and centralized/de- centralized switching for SWI	Uniform random, bit reversal, shuffle, transpose, hotspot, butterfly, table-based, broadcast 1-to-all and multicast 1-to-M	wire, SWI	performance and power

### 3.1 INTRODUCTION

As a result of technology scaling, the number of integrated IPs inside a single SoC has increased dramatically and this has led designers and researchers to adopt the NoC as the underlying communication structure [2, 4, 5, 6]. However, the size and volume of inter-IP communication and the number of IPs are still continuously increasing and causing a growing burden on the NoC. Thus, the regular metal-based NoC (that transmits the signal by charging/discharging the whole RC wire) struggles to match this scalability in terms of latency and energy (J/bit) [7, 8]. This due to the fact that the cross-section of the wire is decreasing [7], and thus wire resistivity is increasing which causes higher power dissipation. Therefore, the issue is not only that the metal wire does not scale enough to match future interconnect requirements, but also that the situation is projected to get worse in terms of power and performance. In other words, we do not have Moore's law for interconnects.

A typical solution is to introduce repeaters for global and semi-global wires, which has been shown to mitigate the delay problem by making it increase linearly with distance. However, this increases the power consumption and area overhead. Therefore, much work is being done on managing repeaters placement and minimizing their numbers and size with an acceptable delay penalty [69]. The drawback of such solutions is that neither power consumption nor latency are optimal.

A more promising way is to look for alternative communication fabrics such as RF-based [24, 82, 18, 19] and optical interconnects [27, 28, 130]. These also seems to be far from ideal, as discussed in Section. 2.3.1. The Zenneck surface wave interconnect is one of the emerging types that might be the optimal solution for satisfying on-chip communication issues. This technology involves an electromagnetic wave that propagates and is guided through an interface between different media surfaces. This chapter investigates this proposed alternative interconnect in terms of performance, power dissipation, and reliability. The major contributions of this chapter are as follows:

- A SWI link design and integration requirements are presented that provide the targeted bandwidth and performance to deliver a flit across the chip in one hop.

- An analytical model is proposed for the SWI based on experimental results.
- A set of experiments are conducted to confirm and establish main concepts regarding the surface-wave technology realisation and features.

### 3.2 SURFACE-WAVE INTERCONNECT FABRIC

#### 3.2.1 *Surface Wave Interconnects (SWI) implementation*

This section discusses the requirements for implementing surface wave for on-chip interconnects. The Zenneck surface wave is an inhomogeneous plane wave supported by a surface. Zenneck surface waves require a surface to have appropriate surface impedance for it to propagate along. These types of surfaces can be realised in a number of ways, however only inductive surface impedances exist in nature, as mentioned in Section 2.3.1.5. Moreover, maximum transmission into the Zenneck surface wave occurs when the incoming wave is incident at or close to the Brewster angle, where reflections are minimized. According to Hendry and et. al. [105], the E-field decay rate in Zenneck surface wave from the source horizontally along the boundary should be around  $(1/\sqrt{d})$ , where  $d$  is the distance from the source. On the other hand, the decay is exponential vertically away from the boundary. This allows less power dissipation for far larger coverage areas than the regular wireless RF, as mentioned in Section 2.3.1.5.

The designed surface media consists of a dielectric layer placed over a corrugated conductor layer [106, 105, 107]. The surface can be engineered by altering its dimensions and materials of conductor and/or dielectric used so that the characteristic impedance ( $Z_s$ ) will be around  $(10 + j300) \Omega$ . In this thesis, for low fabrication costs and simple geometry, the flat surface of dielectric-coated metal sheet is preferable and will be considered. According to Equation 2.3, the waveguide surface can be realised using either silicon dioxide ( $\text{SiO}_2$ ,  $\epsilon_r = 3.9$ ) or ceramic ( $\text{Al}_2\text{O}_3$ ,  $\epsilon_r = 9.8$ ), on a metal ground plane of thickness  $1\mu\text{m}$  ( $\sigma = 3.5$ ). In the case of millimetre-wave applications at 60 GHz or above, the thickness of the dielectric layer for silicon dioxide and ceramic will be 0.8mm and 0.7mm respectively, and the coating process can be integrated with a conventional semi-conductor fabrication process. In addition, due to the fact that the surface roughness of the dielectric layer will not be an issue when the operating frequency of the system is less than 300 GHz, no expensive highly polished wafer ( $R_a < 0.01\mu\text{m}$ ) is needed. As a result, the cost of the additional process can be neglected. On the other hand, in order to dissipate the heat of the chip to the chip heat-sink, the dielectric material should be thermally conductive dielectric. Such dielectric has been widely



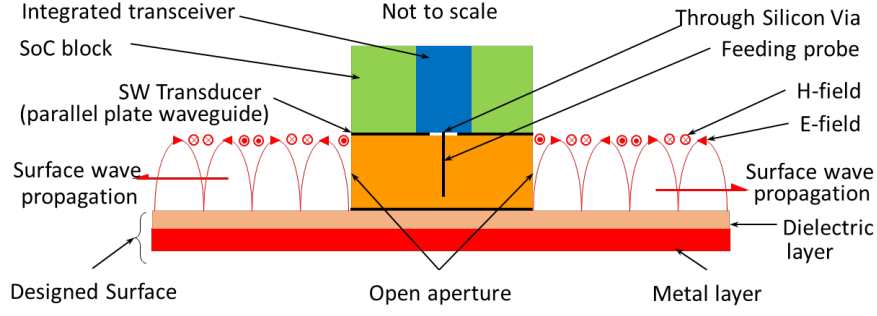


Figure 3.1: Integrated transceiver and integrated transducer (inverted quarter-wavelength monopole) stacked over the designed surface.

investigated and many patents and studies can be found in literature [131, 132].

As mentioned earlier in Section 2.3.1.5, there is no limitation for frequency range carried by the surface. The only limitation, to the authors' best knowledge, is the switching speed of the CMOS technology. Nonetheless, predictions show an increase in transistor switching speed with technology scaling, which would enable the use of higher carrier frequencies [82, 7, 83]. For instance, the maximum carrier frequency for 16nm CMOS technology is 1136 GHz [82]. This range of frequencies is necessary to allow multi-channel realization based on frequency-division-multiplexable-access (FDMA) or code-division-multiplexable-access (CDMA) at this shared media with necessary frequency spacing to avoid channel interference, see section 3.2.2. For this work, a transceiver with FDMA designed for wave guided signal is chosen similar to the transceivers proposed by authors in [19, 98].

On the other hand, an integration of a transducer is needed to launch the waved signal in to the surface. There are a few different ways to implement the transducers for surface communication. It can be as simple as, for directional transmission, a coaxial to waveguide flange as described in [105]. Also, it could be a dipole or monopole for, Omni-directional communication, with parallel plate waveguide [109]. In 3D-EM simulation model shown in Fig.3.1 we used an inverted quarter-wavelength monopole. The transducer layer can be fabricated separately and using flip-chip bonding and TSV technique for connection to the integrated transceiver. The transducer and the transceiver design definitely need further investigation, but it is out of this thesis's scope.

### 3.2.2 SWI Links design

The SWI channels need to be designed to match the requirements of performance critical SoC. The baseline architecture interconnects are designed so that a flit is transmitted through the channel in one cycle. Therefore, the surface wave channels should provide one cycle across

the chip transaction. The cycle period is 0.5ns (frequency 2 GHz) and given that the surface-waves propagate at a speed close to the speed of light, the maximum distance that the signal may travel in one cycle should be less than 15cm. This distance is enough for on-chip, and even off-chip, interconnects.

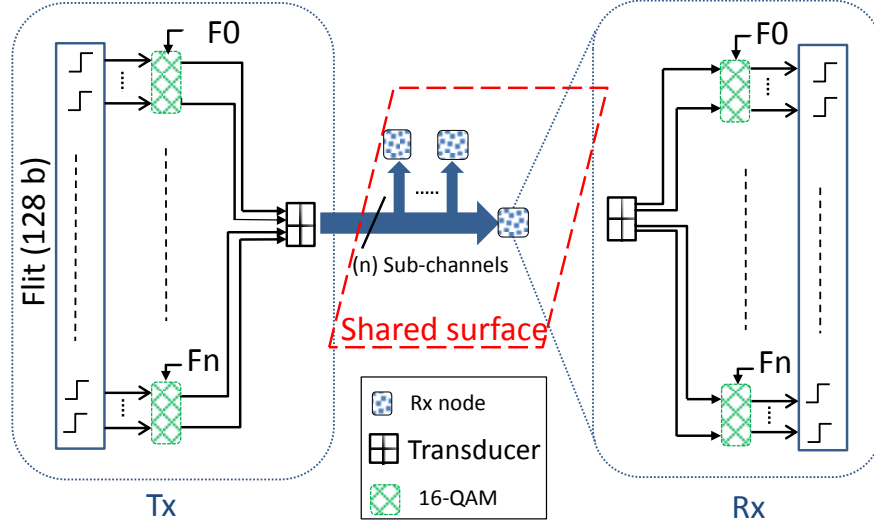


Figure 3.2: Surface wave interconnect communication channel with multi sub-channels where the master node transmit through the shared surface to slave node(s).

On the other hand, the communication channel needs to be designed to handle the same baseline architecture, which is SCC [5], link data rate (256 Gbps) as shown in Section 4.3.1. This can be achieved using an integrated transceiver with capability for multiple access algorithms such as FDMA [98, 101] to give more than one node the right to transmit over the shared surface. In addition, each channel can be designed to have a number of sub-channels (SCs). Each sub-channel (SC) transmits a nibble (4 bits) after it has been modulated using 16-quadrature-amplitude-modulation (QAM) to achieve better bandwidth efficiency (bit/second/Hz) and offer minimum symbol error performance, see section 3.3.2. Furthermore, each SC has a dedicated two GHz frequency bandwidth to guarantee the needed sub-channel frequency spacing. In order to transmit a flit (128 bit) in one cycle, 32 SCs are needed. The total frequency ranges needed per channel (64 GHz) are feasible and even a wider range of frequencies can be achieved in the future as projected [7, 18, 82]. For instance, a 324 GHz oscillator was already demonstrated for the 90nm technology [83]. Notice that a lower frequency bandwidth can be used if the transmitted flit is serialized. The 32-SC option is chosen as shown in Fig. 3.2, because it offers lower latency for the

targeted performance critical systems, even though it increases area overhead as shown in section 4.4.3.

### 3.2.3 SWI Challenges

The SWI is considered to be one of the newest types of interconnect. Therefore, the potential of this technology requires research to tackle various design and implementation challenges of different levels in order for it to be utilized in future NoCs. Firstly, in terms of component integration, the realization of the SWI requires 3D integration techniques to link the transceiver to the transducer, such as TSVs and flip-chip bonding. This technology still faces some technological challenges, as described in section 2.2.8.

Secondly, in terms of communication and RF engineering, careful consideration is required in the design of the integration level of the transceiver, the surface and the transducer. Otherwise, the SWI may pick up noise signals from any CMPs in nearby integrated devices such as power distribution networks and different interconnect components. This interference could affect either the transceiver or the waveguide surface. The impact on the transceiver can be addressed using techniques similar to those in the WiNoC, as mentioned earlier. However, unlike the WiNoC, the SWI requires less SNR due to the fact that it has less signal power dissipation, as will be discussed further in chapter 3. In terms of interference affecting the designed surface, there are two points that strongly mitigate this. The first point is the spacing and isolation between the surface and the integrated circuits. The second is the reflection of any RF signal unless it is incident at or close to the Brewster angle [35, 106].

## 3.3 ZENNECK SURFACE WAVE MODELLING

This section develops a mathematical model for power dissipation and discuss the physical link performance and reliability, which will be used in system level evaluation sections in the following chapters.

### 3.3.1 Analysis of link power dissipation

This section provides the mathematical modelling for the measured power dissipation in Zenneck surface-wave for different frequencies and different distances. Fig.3.3 shows the measured voltage gain ( $S_{21}$ ) in terms of distance for wide frequency range on the designed surface with reactance  $X_s = j188.5$ . The proposed surface acts as waveguide of the propagated signal, with unique characteristic impedance that offers a better attenuation rate than a transmission line. Nonetheless, the transmission line model can be adopted for the modelling of power

dissipation of surface-wave. Thus, Equation.3.1 that represents the transmission line voltage decay equation with distance fits perfectly:

$$|V^+|_d = |V^+|_0 e^{-\alpha d} \quad (3.1)$$

Thus:

$$\frac{|V^+|_d}{|V^+|_0} = e^{-\alpha d} \quad (3.2)$$

Therefore, the voltage gain in dB is:

$$S_{21} = E + 20 \log_{10} e^{-\alpha d} \quad (3.3)$$

Where  $S_{21}$  is the signal forward voltage gain in dB,  $|V^+|$  is the amplitude of travelling wave,  $\alpha$  is the attenuation constant (dB/m), and  $d$  is the propagation distance (m).  $E$  is the loss constant (dB) resulted from the inefficiency of signal transformation between the surface waveguide and transceiver due to the use of an off-the-shelf transducer in the experiment set-up [106] where the surface was designed to resonate at 23GHz. The Matlab curve fitting tool is applied to extract the value of  $\alpha$  and  $E$ . This procedure was repeated for a set of frequencies ranging from 23-35 GHz. The calculated attenuation constant average is 6.33 neper/m and average  $E$  is -23.8 dB. Fig.3.4 shows calculated versus measured voltage gains with an error average percentage at around 1.15%. These values are used in simulations to calculate the link power dissipation in this thesis.

### 3.3.2 Communication system performance

It is important to discuss and identify SWI channel capability and reliability at this point. The chosen sub-channel bandwidth is 2 GHz to fit the required data throughput to avoid inter-symbol interference, thus minimizing the error rate. The targeted channel BER is less than  $10^{-14}$ . Moreover, each channel consists of 32 sub-channels with 16-QAM signal modulation. Thus, the calculated SNR needs to be more than 19 dB for this communication system as shown in Fig.3.5.

On the other hand, for a SoC with many integrated devices, a power distribution network and/or different interconnect fabric could inject noise signal either to the the SWI or to the transceiver. However, as mentioned in Section 2.3.1.5, giving the right spacing between the surface and integrated circuits and the maturity of integrated RF devices especially the transceivers will significantly mitigated the interference. However, for a reasonable assumption we should consider that the channel is not noise free. As presented in [133, 98] a thermal limit ( $Th_N$ ) is used to set the noise floor (-174 dB), SNR determined to

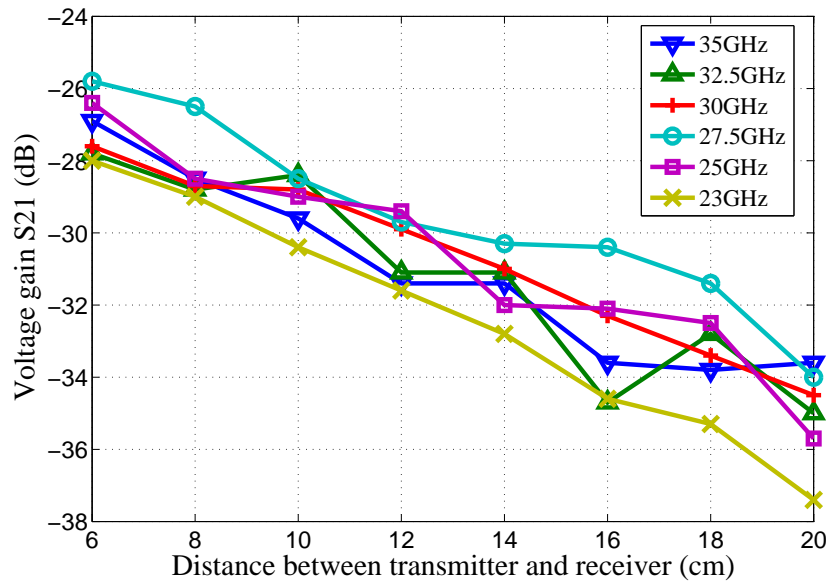


Figure 3.3: Experiment results showing power decay with distance for a range of frequencies [35].

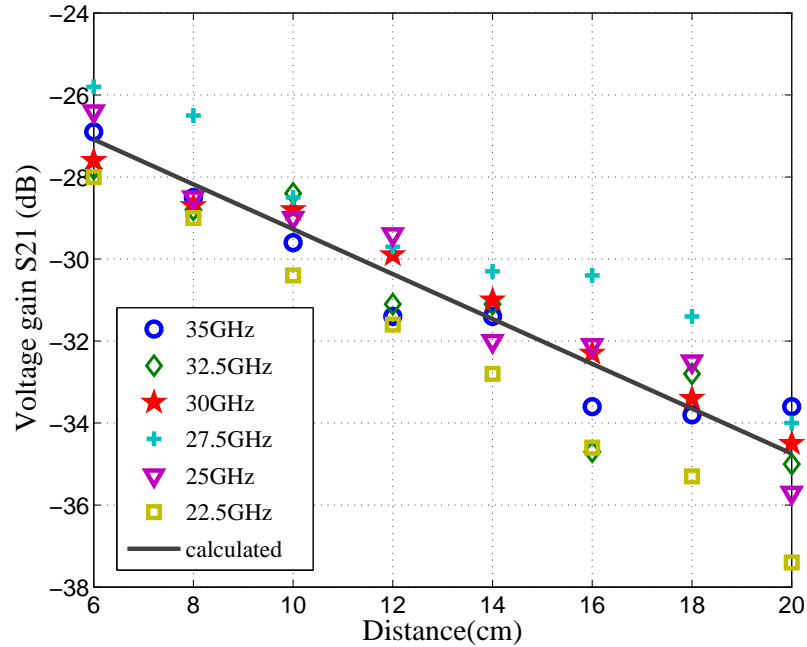


Figure 3.4: Measurement versus calculated voltage gain comparison for different frequencies.

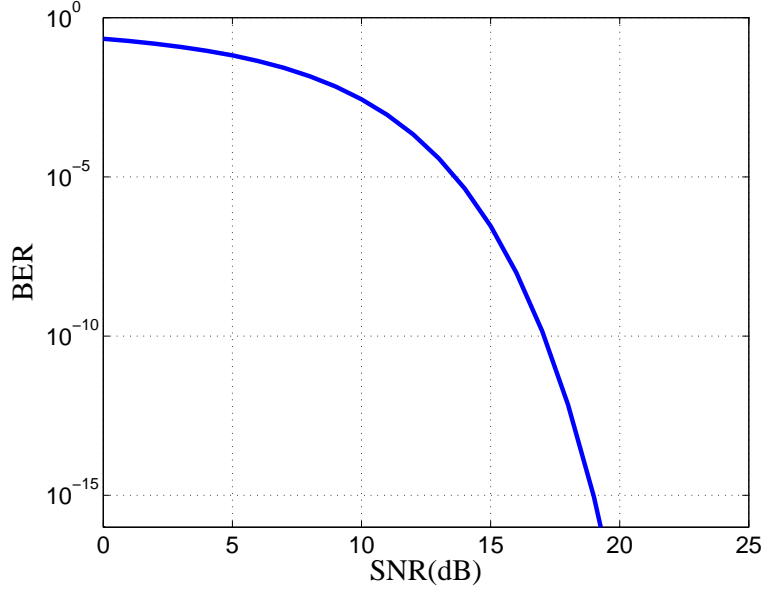


Figure 3.5: Bit-Error-Rate vs. SNR for 16-QAM modulation sub-channel (SC).

be (20 dB), and noise figure for the receiver ( $R_{xN}$ ) assumed to be (3 dB). These limit the minimum error free received power per sub-channel to:

$$P_{\min} = 10 \log_{10} (\Delta F) + \text{SNR} + R_{xN} + T_{hN}, \quad (3.4)$$

Where  $\Delta F$  is the frequency bandwidth for one channel. And for a channel bandwidth of 64GHz the minimum detectable signal is (-43 dBm). Moreover, we will assume a loss of inefficient signal power transaction between the surface and the transducer, which was observed to be  $E = -10$  (dBm), assuming that a better designed integrated transducer is being used, unlike the transducer used in real experiments, see section 3.3.1. Therefore, Equation(3.4) can be used to calculate the surface wave decay rate per distance. Consequently, the minimum received power should satisfy this equation:

$$0 > |E + 20 \log_{10} e^{-\alpha d}| + P_{\min}, \quad (3.5)$$

Based on this equation, a signal is valid for a distance close to 60cm with this low BER constraint. This distance is suitable for on-chip and on-board communication. A deeper and more detailed analysis and investigation are needed to the communication system, transducer design and the noise figures. However, this is out of this thesis scope. Nonetheless, this level of analysis is serving our purpose since it gives first order estimations that can be used by SoC designers to assess this technology for on-chip interconnects fairly on a system-level abstract.

### 3.4 EXPERIMENT RESULTS

This section presents experiments that aim to confirm and/or establish a set of main concepts regarding surface-wave communication. Firstly, test the hypothesis that surface-wave is better than free-space electromagnetic propagation in terms of power dissipation. Secondly, confirm previous experiments and simulations conducted in related work. Thirdly, compare between the corrugated and the flat surface. Finally, prove that surface-wave can be realized even when off-the-shelf equipment is used due to limited funding for instance.

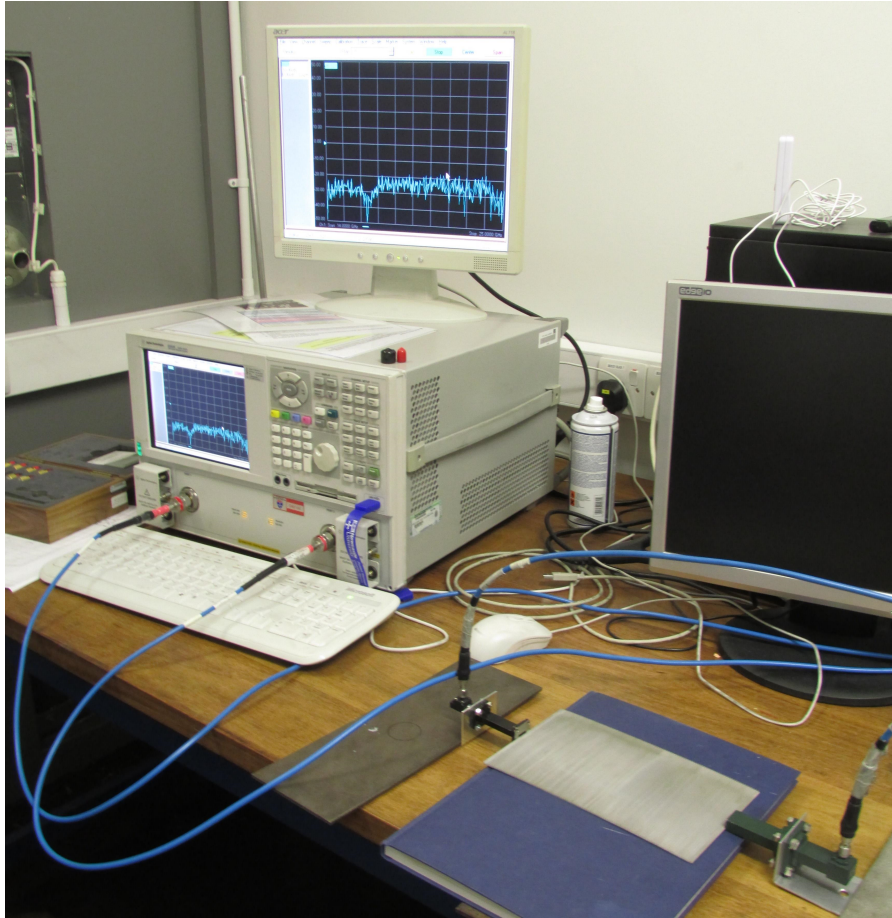


Figure 3.6: The surface-wave experiment set-up.

Fig. 3.6 shows the experiment set-up for the corrugated surface. In this set-up, two different coaxial to waveguide adaptors have been used. These two transducers, according to their data-sheet and experimentations have a common frequency range between 14 and 25 GHz. These transducers then were connected to a vector network analyser (VNA) type Agilent-8363B. The transducers lowest aperture point is the designed surfaces top edge where the dielectric materials in the surface is to the top. Calibration for this wide range and with the available equipment and connectors has been proven to be not ideal.

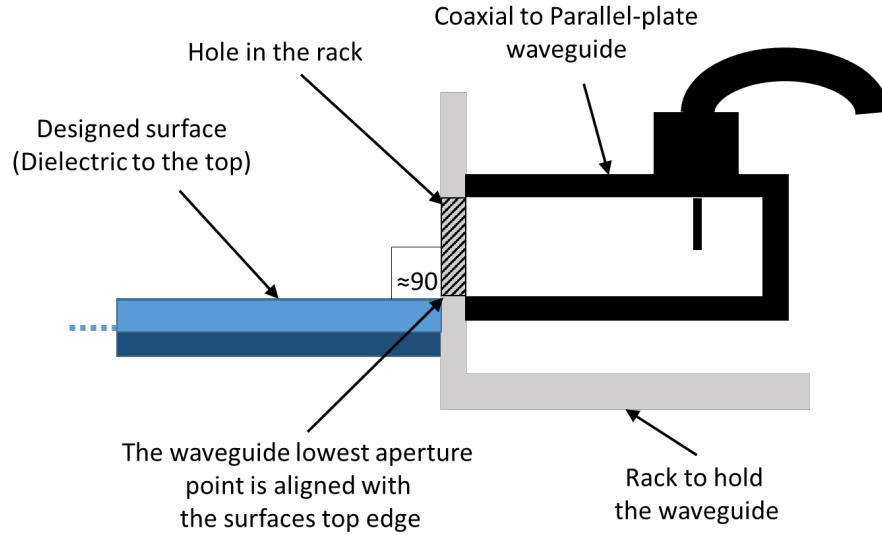


Figure 3.7: The surface-wave experiment equipment alignment and set-up.

This is due to either the imperfection of equipment alignment, such as the transducer angel to the surface or the alignment of the transducer edge as shown in Fig. 3.7, or the mismatch between the used transducer/waveguides. This is clearly shown in Fig. 3.8 where waveguides are connected face-to-face without any intermediate medium and yet the  $S_{21}$  measured on average equals to -14.42 dB.

Regardless of all the factors mentioned above, experiments show improvements in  $S_{21}$  up to 20 dB and 10 dB on average over free-space propagation when a corrugated surface has been inserted between the two transducers for a distance of 20cm. This corrugated surface grooves have dimensions which are  $d_1 = 2\text{mm}$ ,  $d_2 = d_3 = 1\text{mm}$ , see Section 2.3.1.5. The second designed surface involve a two off-shelf sheets of polyethylene terephthalate glycol (PETG) of 1.2 mm thickness and flat aluminium sheet of 1mm thickness. The latter surface shows even better signal propagation with an average of  $S_{21}$  equal to -21.04 dB, which is higher by 4 dB than the corrugated surface.

Fig. 3.9 shows the attenuation of the signal in distance as the distance between the transducers changed from 20, 10, and 6 cm for the flat surface with average  $S_{21}$  -21.04, -18.7, and -16.49 dB, respectively. Thus, according to Equation 3.3 the calculated average attenuation constant ( $\alpha$ ) is around 4.8 neper/m. Although, these results confirm the reported attenuation relation with the propagation distance, they have not been used in our analytical power dissipation model since the experimentation set-up was not optimum as mentioned earlier. Nonetheless, the experiments successfully achieved their goals in validating a set of main hypotheses presented in this chapter that are considered to be the ground for system-level evaluation in the following chapters.



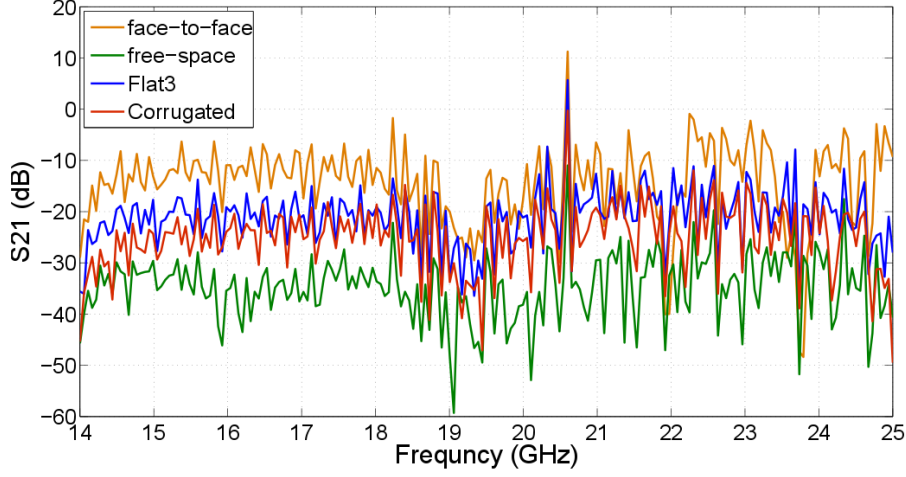


Figure 3.8: S21 measurements results for the corrugated surface, flat surface, free-space, and the case where transducers connected face-to-face for wide range of frequencies.

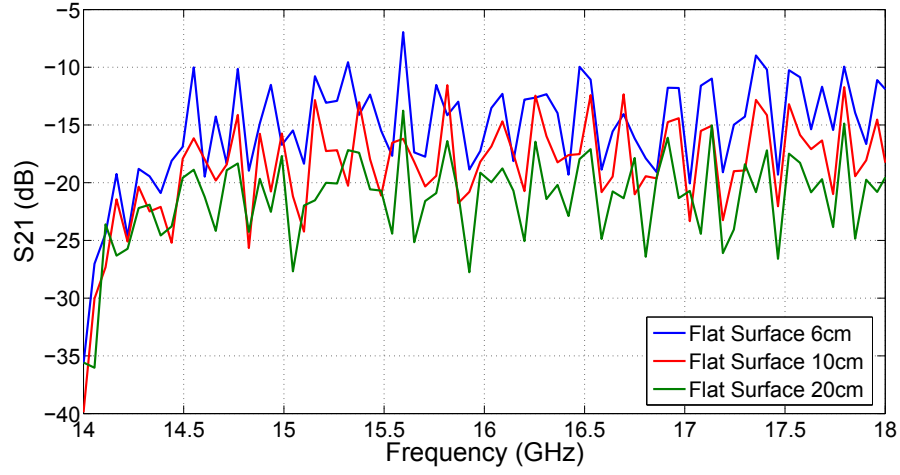


Figure 3.9: S21 measurements results for the flat surface shows signal attenuation with distances for wide range of frequencies.

### 3.5 SUMMARY AND CONCLUSION

This Chapter aims to propose and discuss SWI as future promising on-chip interconnect. Zenneck surface wave low power dissipation, CMOS compatibility and high signal propagation speed might allow significant mitigation of intra-chip communication issues that are projected to scale up as the technology scales down. Implementation considerations for this interconnect have been discussed such as the needed RF integrated devices and communication channel design. In addition, the SWI system level power modelling and performance considerations have been discussed. System level evaluation in the rest of the thesis will be based on this analytic model and considerations. Moreover, experiments on different types of designed surface

compared to free-space propagation to confirm analytical power and performance model. Consequently, experimentation, analysis, and modelling would allow the study of potential improvements and challenges for any SWI-based communication architecture.

## HYBRID WIRE AND SURFACE-WAVE ARCHITECTURE FOR ON-CHIP GLOBAL COMMUNICATIONS

---

### 4.1 INTRODUCTION

Integrated circuit (IC) technology processes are scaling rapidly due to growing market demands for more complex SoC. This is leading to an intensification in terms of IPs/core density and functional complexity of current and future SoC. Moreover, the inter-IPs/core communication increasingly becoming a major factor in determining performance and power consumption of SoC. This is due to the fact that CMOS technology scaling is not in favour of wires. This challenge is only mitigated by a major shift in the designing of on-chip interconnect architectures from the bus to the NoC architecture where such an architecture consists of a network constructed from multiple point-to-point data channels (links) interconnected by routers, in other words multiple hops. This is due to the fundamental wire issues especially for global cross-the-chip communication and the increasing of the number of intermediate hops, and therefore its delay, for such architectures. In contrast, for local communication, since wire length is scaling down with the gate size, the wire seems to be still efficient. This is different from global communication in which the wire length remains the same yet the cross section of the wire is decreasing [7], and thus wire resistance is increasing, which causes higher power dissipation. Consequently, the regular metal-based NoC will no longer match this scalability in terms of latency and energy (J/b) [7, 8] and new interconnect architectures are required.

Novel solutions include the use of alternative communication fabrics such as radio RF-I [24, 82, 18, 19] and optical interconnects [27, 28, 130] as supplementary interconnects for global communication. These interconnects seem to have some drawbacks that yet have to be addressed, as discussed in section 2.3.1. Zenneck surface wave might be able to mitigate global communication issues due to its remarkable features in terms of power efficiency and CMOS-compatibility. This chapter proposes a new architecture that utilizes merits offered by this proposed alternative fabric to mitigate the projected increasing global communication and multi-hop issues. The major contributions of this chapter are as follows:

- A hybrid wire and surface-wave interconnects architecture utilizing the Zenneck surface-wave features is introduced to mitigate global communication issues.

- A distance-based weighted-random arbitration algorithm is proposed that intelligently prioritizes surface wave channels over metal wires in the case of global communication.
- Results show improvements in average delay (34%) and throughput (35%) and power consumption (12 to 23%) over low cost in terms of area overhead.

## 4.2 RELEVANT WORK ON CROSS-THE-CHIP COMMUNICATIONS

### 4.2.1 *Multi-hop Challenges and Wire-based Solutions*

Despite the fact that NoC is far more scalable than the bus architecture, the packets are required to travel via multi-segments of wires linked by routers. To pass through these routers, in turn, packets need to go through their pipeline which range from 2-4 cycles depending on router's micro-architecture. The travel from one router to another is referred to as a hop. Moreover, the average-hop-count among nodes is a metric that defines the NoC latency and performance, and it is scaling with the number of routers for many NoC topologies such as mesh.

To avoid this issue, Ogras et al. targeted application-specific networks by inserting shortcuts links on a mesh NoC based on studying characteristics of application traffic [134]. This work was then developed further to target the small-world principle, which improves NoC performance by inserting these long-range links. On the other hand, Kumar et al. suggested to use express VCs to reduce the time to bypass routers [135]. A configurable NoC was also suggested to introduce direct or long-range links by skipping intermediate routers. For instance, the ReNoC includes a configurable layer between routers and links that can be customized based on the application [136]. Moreover, configurable multidrop express links have also been proposed [137]. However, all these novel architectures are limited not only by the extra hardware they require but mainly by issues associated with wires in terms of latency, bandwidth, and power dissipation (see Section 2.3.1.1). Therefore, other emerging interconnects have been introduced to overcome these limitations.

### 4.2.2 *Routing in Hybrid Architectures*

Novel routing schemes are needed to forward traffic in hybrid architectures with a multilayer network, irregular topology and different physical interconnect specifications. These routing schemes should aim to maximise the utilization of the under-layer fabric while avoiding deadlocks or hotspots. Wettin et al. developed routing algorithms for irregular networks [138, 139, 140], which offer deadlock-free routing

and alleviate hotspots, and some of them offer flexible routing based on NoC status. However, these routing algorithms have relatively high complexity and area overhead. For instance, both multiple-tree-roots (MROOTs) and adaptive-layered-shortest-path (ALASH) algorithms require a set of dedicated VCs.

Other studies have used routing tables in each router [18], which offer reconfigurability to the NoC. However, these routing techniques also involve extra area overheads for the routing tables that scale exponentially as NoC size scales. Moreover, if the NoC needs to be reconfigured, many cycles might be required to update all of the distributed routing tables.

Therefore, for many state-of-the-art hybrid architectures, simple turn models have been proposed to route packets via the wire-based layer, which are usually connected as a mesh. Then to route packets via the second NoC layer such as RF-I [82] and WiNoC [103, 94], a few modifications or special routing schemes are used. These routing schemes involve low area overheads to the router micro-architectures. However, careful consideration is required with wormhole flow control in order to avoid deadlocks, since the express links added to the wire-based mesh might create cyclic dependency.

#### 4.3 HYBRID WIRE AND SURFACE-WAVE INTERCONNECT ARCHITECTURE

Hybrid interconnect architectures that combine the shared medium bus and the indirect network have been previously shown to have limited scalability in performance-critical SoCs due to the limitations of the wire-based bus layer [3]. Hybrid networks could retain many merits of the buses such as the 1-to-M capability and reducing inter-node average hop count while maintaining high interconnect scalability if a high performance interconnect was adopted as the bus network layer [3]. Therefore, the use of the SWI as a supplementary interconnect, with the favourable merits is proposed here. This section presents schemes to maximize the utilization of W-SWI for global cross the chip communication.

##### 4.3.1 Addressing Multi-hop Challenges Using W-SWI Architecture

The SWI has significant advantages over many other interconnect technologies, as mentioned earlier. However, as with all RF-based interconnects, it suffers from limitations in terms of shared media and limited ranges of frequencies. These make it infeasible to completely replace metal wire interconnects in the near future [141]. Moreover, in terms of wire-based interconnects, local communication seems to scale well with technology scaling unlike global communication [7].

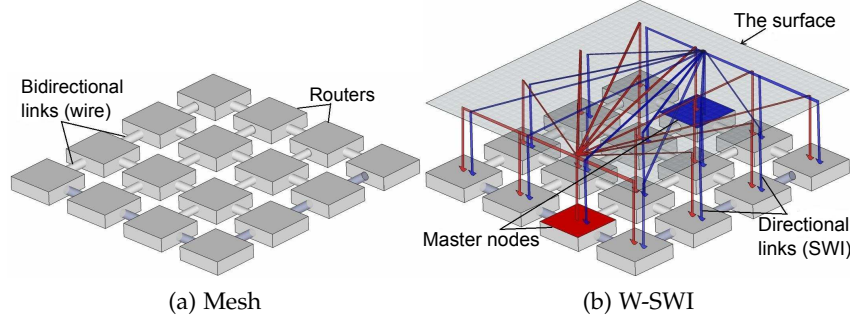


Figure 4.1: Example showing that inserting two SWI channels in the proposed hybrid wire-SWI multilayer-network increases the overall NoC bisection bandwidth: (a) conventional on-chip network layer with 4-ary 2-mesh topology; (b) connections of both layers, metal wire and SWI.

In addition, this type of interconnect has the cheapest implementation cost compared to other fabrics.

Therefore, we argue that the best solution would be to combine both metal and surface-wave interconnect in hybrid wire-SW interconnects to create a multi-layered network architecture; in this thesis it will be referred to as W-SWI, as shown in Fig. 4.1. The first layer is a regular mesh topology, which is preferable for general purpose interconnects. The second layer is the surface wave bus topology. Thus, this architecture offers a natural fanout feature, which is lost when the interconnects system changes from the bus to the NoC. Moreover, it substantially increases the network bisection by  $N_m(N - 1)$ , where  $N$  is the NoC size and  $N_m$  is the number of number of nodes with transmission capability via SWI channels. Therefore, the reliability of the overall system is improved by increasing the network bisection, which makes it more robust to any type of failure that may isolate part of the NoC.

In addition, it can be seen that the extra links bring the NoC closer to achieving the small-world phenomena [142] where every node have access to other nodes in one hop. Fig. 4.2 shows the effect of introducing and increasing  $N_m$  on the average and maximum hop count assuming minimum rout. Clearly, even small number of  $N_m$  has significant effect on closing the gap between the resulted topology and small world. As a result, even small number of SWI would increase the NoC capacity to higher PIR.

In order to preserve the fanout feature, all the routers in the NoC are designed to receive information through SWI. On the other hand, if all the nodes have transmission capability, this will increase connectivity. However, this would increase the contention on the SWI layer as the NoC size increases. In addition, multicast communication and global communication are relatively low but with a dramatic effect on NoCs performance, as mentioned will be discussed in the next chapters.

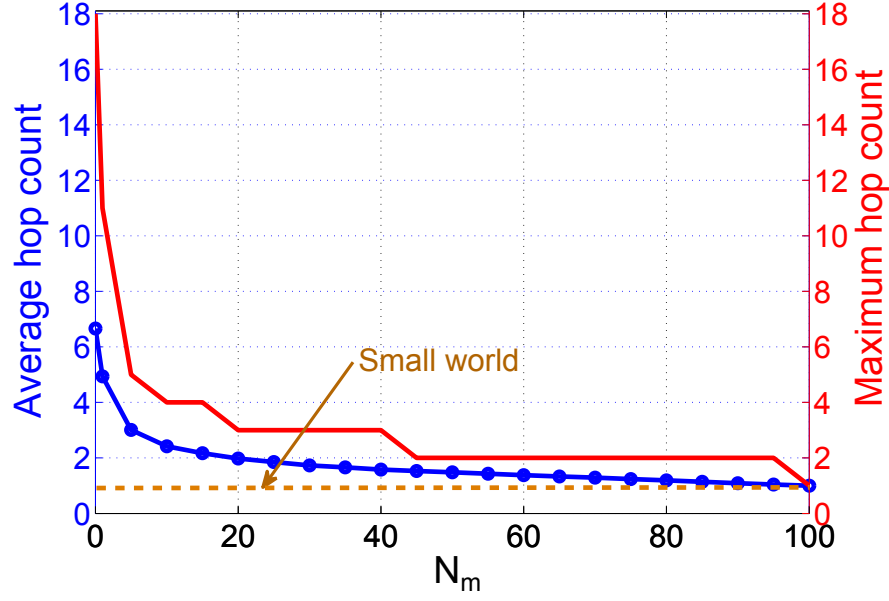


Figure 4.2: Closing the gap towards small world phenomena in a  $10 \times 10$  NoC as the number of nodes with transmission capability via SWI ( $N_m$ ) increased.

Moreover, as mentioned earlier small number of  $N_m$  will have significant topological advantage in terms of hop count. Therefore, fewer nodes are selected to have the transmission capability to reduce the circuit overhead and comply with the available frequency bandwidth. These nodes will be referred to as masters, while the rest are referred to as slaves. These slaves can only receive data but they may transmit some control signals, see chapters 5 and 6.

Masters can be distributed based on many targets such as minimizing the maximum-hop-counts in the NoC or average distance between the nodes [143]. Moreover, master locations could be determined using optimization algorithms for general or specific application patterns to achieve optimum performance and/or power consumption [144, 145]. However, in this work, masters are distributed so that the average hop count (Manhattan distance) from all slaves to the nearest master is at a minimum. This placement of the master nodes reduces the average hop count of packets via the costly wires and routers in the network. In addition, it would allow each master to be the centre of a sub-network region. As a result, the masters would be accessible with minimum hop count via wires and routers for critical traffic such as the 1-to-M or global communication, thus mitigating their impact, which is the target of this thesis. In this study, The master placements calculations have been determined using simple simulated-annealing optimization algorithm targeting lowest average-hop-count from all slaves to the nearest master. Fig. 4.3 shows an example of such placement of 4 masters in  $6 \times 4$  NoC.

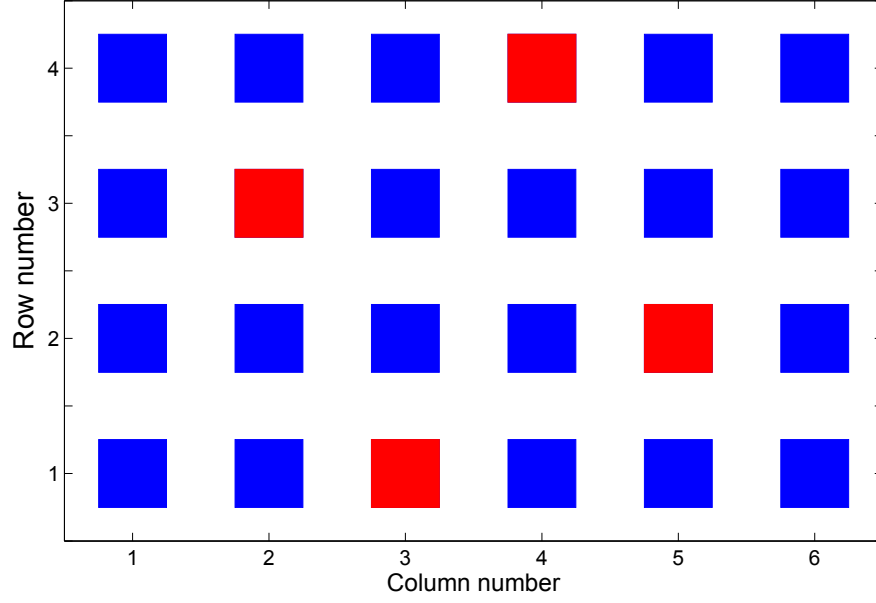


Figure 4.3: Example of 4 master placement in a  $6 \times 4$  NoC based on simple simulated-annealing optimization algorithm with target to minimize the average-hop-count from all slaves to the nearest master.

In this thesis SCC [5] is adopted as the baseline architecture (Baseline-Arch). This chip is designed for performance critical multicore processor, which makes it optimum for this work aim. However, a sixth port needs to be added to the router with all related control circuits. Also, a crossbar switch size needs to be adjusted according to the newly added port. This port is linked to a transceiver with either transceiver (Tx/Rx) capability or just receiver (Rx). The tile specifications and parameters, which are used in all the thesis simulations and results sections, are shown in Table 4.1.

#### 4.3.2 Routing scheme

Wormhole flow control is used in the hybrid W-SWI architecture because of its low latency and small routers buffer size required compared to other cut-through switching [51, 3]. As mentioned earlier, in hybrid architectures with two interconnects fabrics used, decision problems rise to route the packet via this irregular NoC topology and non-uniform interconnects fabrics. This routing decision problem should aim to utilize the under-layer interconnects physical fabric while avoiding creating hotspots or deadlocks.

In this thesis, for router micro-architecture simplicity, an odd-even turn model routing is being used for routing in the Mesh [55]. This turn model has been chosen due to its high path diversity compared to other turn models routing algorithms. In order to utilize the SWI express links without creating hotspots on master nodes, there are no



Table 4.1: The adopted parameters of the NoC based multi-processor chip (consists of  $6 \times 4$  tiles).

Each Tile consists of		
IP components	NoC components	
Two Pentium™ class IA-32 cores	Message passing router	4-port to neighbour routers, 1-port to local cores and 1-port for the SWI
		6 buffers each with 3-flit depth and 16 Byte width
Two 256 KB private L2 caches	Links	5 bidirectional wire interconnects (16 Byte width)
		1 surface wave channel (Rx or Tx/Rx)
Tile dimension $3.6 \times 5.2 \text{ mm}^2$		

predetermination, static or dynamic, for packets to be routed via SWI for one-to-one communication and avoid deadlocks. The following routing rules are suggested:

- *Rule 1:* Packet passing via master node may transfer via SWI interconnects since there are no predetermination, static or dynamic, for packets to be routed via SWI for 1-to-1 communication.
- *Rule 2:* Packets are not allowed to be routed again via the Mesh and should be drained directly in the local PE.

These rules can be demonstrated in the example in Fig 4.4. The first rule aims to avoid creating hotspots and congestion in SWI links. Moreover, Section 4.3.3 tries to prioritize the use of SWI by directing global communication via SWI and therefore increase its utilization. On the other hand, since the odd-even routing via the mesh is deadlock free, the only remaining concern is cycle dependency due to routing via the NoC second layer, which is the SWI. The second rule guarantees a unicast deadlock free routing in the proposed hybrid architecture. This theory can be proved by contradiction. Assume packets are creating a resource cycle dependency that include one packet that is routed through SWI. Since this packet is directed to the destination node only where it should be drained locally, then it cannot be waiting for any other packets that aim to travel via the mesh layer. This would break any possible cycle dependency and therefore prevent any deadlock scenario.

#### 4.3.3 Distance-based Weighted-round-robin Arbitration (DWA) Algorithm

This algorithm aims to maximize the gain of the proposed hybrid architecture while avoiding creating a bottleneck. It can be described

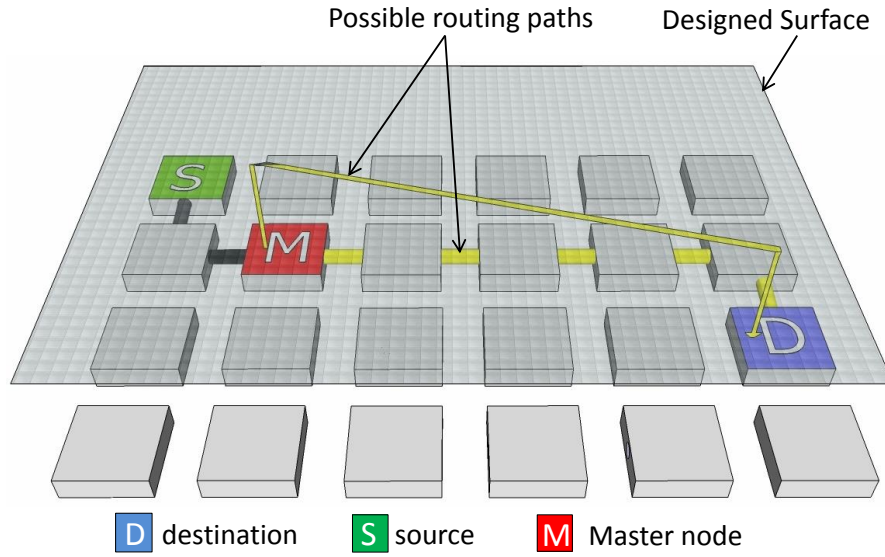


Figure 4.4: An illustration example showing the master node routing decision, which is either forwarding the traffic via SWI or continue via regular Mesh.

as following: First, the possible set of paths for each propagating flit is given by a routing algorithm in each router. Second, a selection operation is carried out by the arbitration algorithm based on the calculated weight ( $W$ ). The proposed Distance-based weighted-round-robin arbitration (DWA) algorithm takes into consideration the distance between the current router and the destination router before selecting surface wave channel as this would increase the power saving and/or enhance network performance (throughput and delay). Therefore, the DWA algorithm (see Algorithm.4.1) is used to give a higher priority to the surface wave channel as the distance increases. Moreover, based on our power analysis presented in Section 4.4.2, we found that if the destination is at close proximity of less than 2 hops then the power cost of the SWI will exceed the power cost of metal links and routers (assuming not loaded path). Figure 4.5 compares the power consumption per distance of SWI and regular wire with repeaters. The DWA should keep the SWI likely available for the flits need to travel long distances (in terms of hop counts) and increase power saving in the NoCs.

This algorithm assigns zero weight ( $W_0$ ) for a distance of one hop and certain start up weight ( $W_1$ ) for 2 hops and increases weight linearly until it gives the maximum weight (100%) for the maximum possible distance (i.e., in case of mesh NoC this distance will be equal to half the NoC diameter).

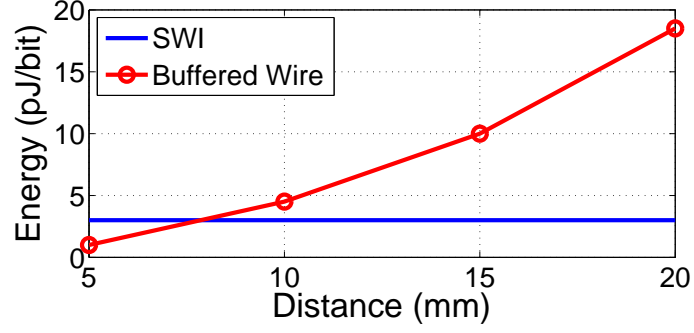


Figure 4.5: Energy dissipation per bit according to on-chip communication distance for buffered wire [22] and SWI.

---

**Algorithm 4.1** Distance based weighted random arbitration algorithm(DWA).

---

- 1: **Input:** NoC Dimension in direction X (X), NoC Dimension in direction Y (Y), distance in hops (d) , P = set of possible Output port from a routing algorithm.
  - 2: **Output:** Chosen output port
  - 3: **if** (d > 1) **then**
  - 4:    $W = (\frac{W_1}{(X+Y)-2} \times (d - 2)) + (100 - W_1)$
  - 5: **else**
  - 6:    $W = W_0$ ;
  - 7: **end if** Circular Shift Right(W);
  - 8: **if** (W[LSB] == 0) **then**
  - 9:   **Return** : Surface wave channel;
  - 10: **else**
  - 11:   **Return** : C, where  $C \in P$ ;
  - 12: **end if**
-

To implement the proposed DWA algorithm, we can use a circular-shift-register (CSR) to hold a code that will represent  $W$ , see Fig.4.6. This weight is calculated and stored in each master node in advance, such as design time or at NoC configuring time. For example, to give SWI a weight of 70% (metal wire weight is 30%) we can store the code (0010010010) in CSR, where if the value of the least significant bit (LSB) is (0), it will forward only SWI signal generated from the used routing algorithm. Otherwise, if LSB value is (1), it will forward the signals of the regular ports linked to metal wires. The CSR will be shifted after each access. This is to do with how the Round-robin algorithm works. The decoded flit destination is the address where the right weight word stored in CSR is accessed. Flit destination is also used as the selector for the multiplexer (see Fig.4.6). The size of CSR is proportional to the weight precision required. On the other hand, the number of CSR is equal to  $(N-1)$ , where  $N$  is the total number of nodes in the NoC. The circuit drawn in solid line in Fig.4.6 is the additional hardware required to implement our proposed DWA. Moreover, this hardware is required only if the router is situated in a master node. Thus the power and area over head is almost negligible, see section 4.4.3 and 4.4.2.

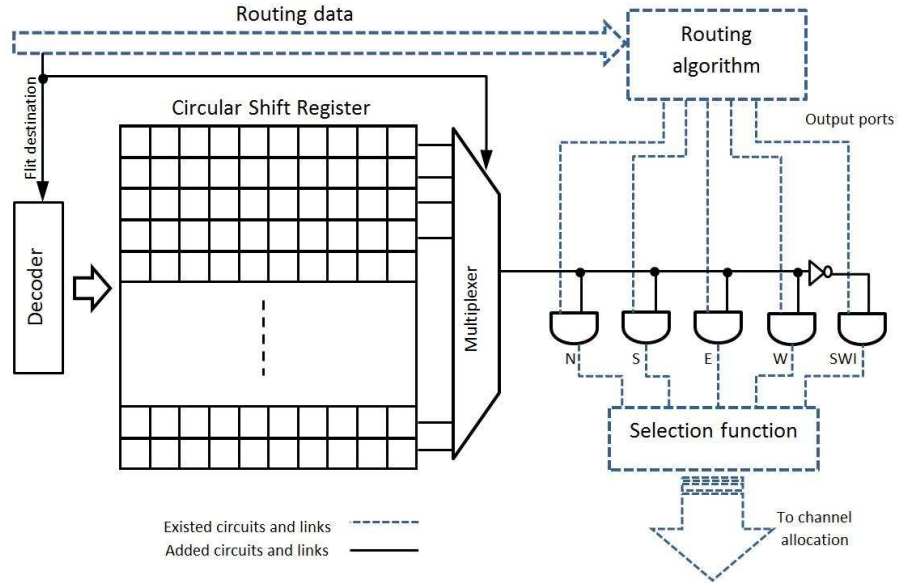


Figure 4.6: DWA implementation, where CSR provide the round robin functionality and store the weight code.

#### 4.4 SYSTEM LEVEL EVALUATION AND DISCUSSION

This section presents results obtained from our cycle-accurate NoC simulator which was built by modifying the existing Noxim simulator [121] for the W-SWI and the baseline architecture of regular mesh; in short Mesh. The Mesh in this chapter refers to a wire-based mesh NoC.

Table 4.2: Scaling of number of SWI channels as the NoC scale based on the assumption that the CMOS technology has scaled.

CMOS technology (nm)	45	32	22	16
Max. carrier Frequency (GHz) [82]	592	768	944	1136
NoC size	6×4	8×8	10×10	12×12
SWI channels	4	5	6	8

In this thesis, the Intel SCC [5] is adopted as the baseline architecture, as shown in Table 4.1. This chip is designed for performance critical CMPs, which makes it optimal for the purpose of this study. The number of master nodes is based on the available frequency range for 45nm, which was estimated to be five channels. However, this frequency range is scaling with technology [18, 82], see Table 4.2. As a result, the number of master nodes was increased when simulating larger NoCs, assuming that the technology will have been scaled, too. Packet sizes 12 flits were chosen as an example to demonstrate the behaviour of the proposed architecture under wormhole-switching flow control. Fig. 4.7 presents a flowchart for the procedure and tools, or models, used to conduct the experimental evaluation in this chapter.

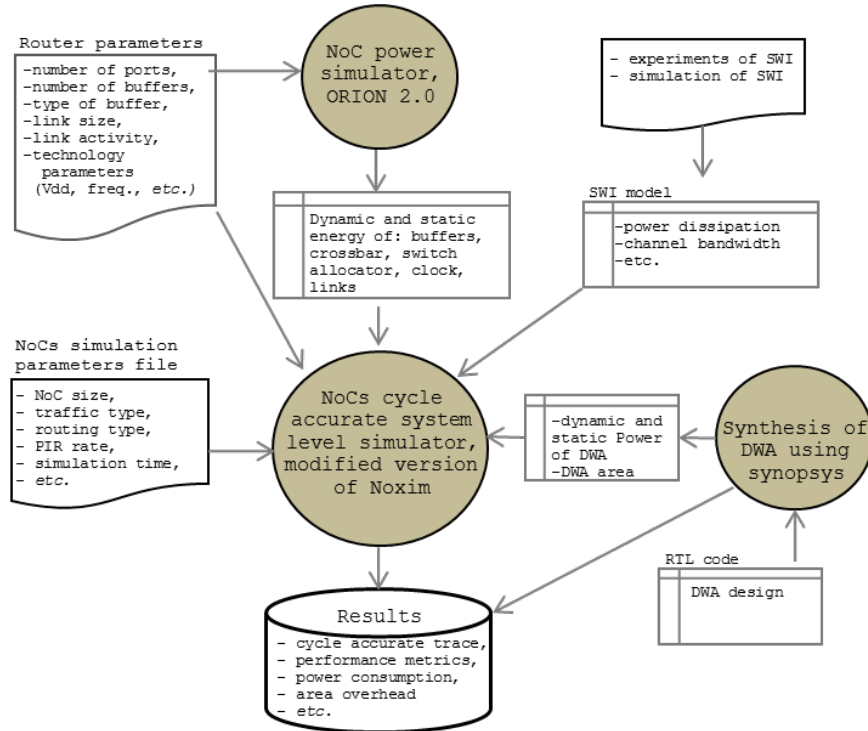


Figure 4.7: The conducted system-level evaluation flowchart shows the methodology and tools used to obtain the results.

The simulation was conducted with synthetic traffics, which are:  
(1) Random, where packets are transmitted randomly with uniform

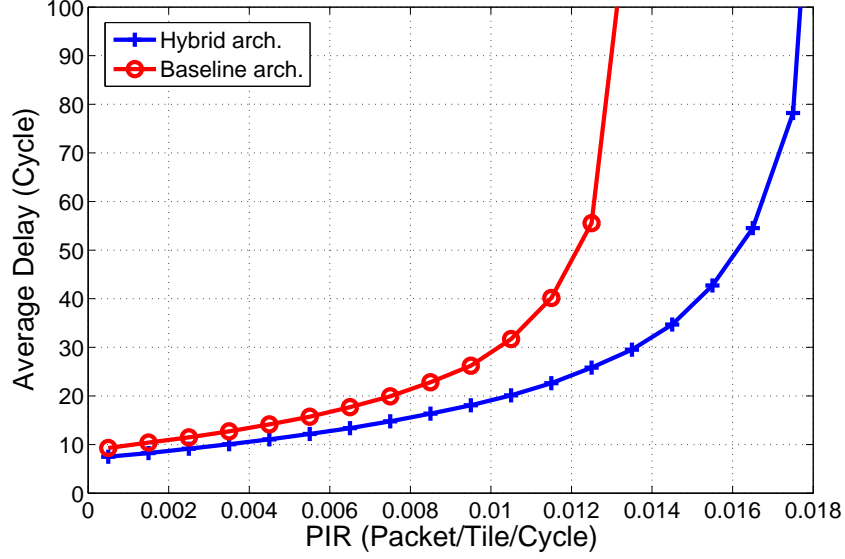


Figure 4.8:  $6 \times 4$  Network average delay verses PIR for W-SWI and baseline architecture.

probability to other nodes; (2) Hotspot, which is the same as the random but with specific nodes called hotspots, four in this case which are either at the edge of the NoC (Hotspot<sub>1</sub>) or the centre (Hotspot<sub>2</sub>), with higher probability of traffic dispatched to them; (3) Transpose, where a node sends a packet to another node that has its address transposed; (4) Butterfly, the destination address swapping the MSB with the LSB of the source; (5) Shuffle, rotate the address one bit; (6) Bitreversal, the destination address is swapping all bits of the source address.

#### 4.4.1 Performance Evaluation

This section presents results obtained from the cycle accurate NoC simulator for the hybrid architecture and compares it with the baseline architecture. The DWA algorithm, proposed in section 4.3.3, was used in route selection for master nodes.

Results had shown significant improvements in performance in terms of average delay. For instance, Fig.4.8 shows less average delay for hybrid architecture over the baseline architecture. This is simply due to the fact that some packets no longer need to travel through routers and costly wire links if they use the SWI highway to reach its destination in one hop. Also, these extra channels are enabling the NoC to handle higher throughput rates than the baseline architecture as shown in Fig.4.9. Table 4.3 shows the use of hybrid interconnect improvement rate over the baseline architecture for different types of traffic. These statistics are taken with limiting the average delay to

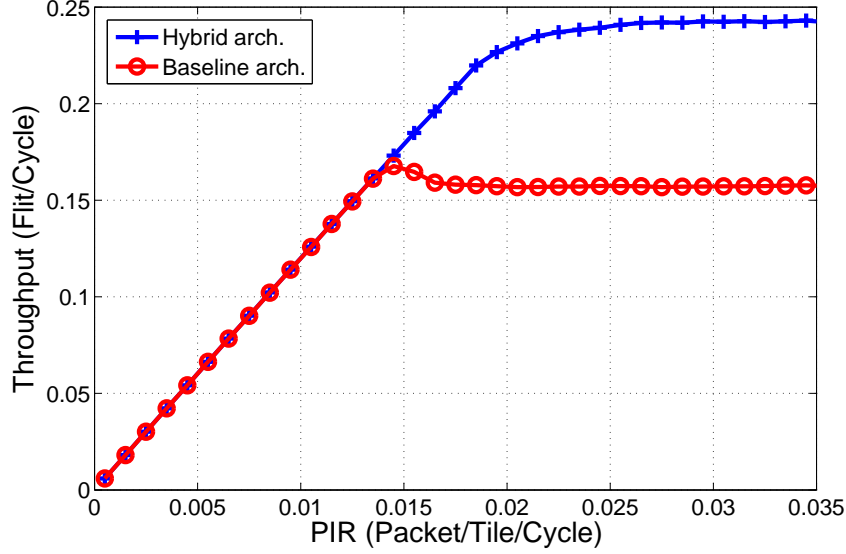


Figure 4.9:  $6 \times 4$  Network throughput verses PIR for W-SWI and baseline architecture.

double the zero-load-latency time [51], which represents the edge of NoC saturation point for fair comparison. The average improvement for PIR and the average throughput is 35% and 34% over the baseline architecture, respectively.

Moreover, to test the proposed architecture on real application benchmarks, we adopted three different size tests used in [146]. These benchmarks are mapped to NoC platforms with the aim of reducing the overall communication energy. Table 4.4 presents the performance improvement gained using the proposed architecture compared to the baseline architecture after running the simulation for 1,000,000 cycles using these benchmarks. The results are consistent with synthetic traffic presented earlier. Moreover, results are shown that as the NoC size increases the global communication issues grow, thus the hybrid wire-SWI architecture has better performance improvements. On the other hand, the energy-aware mapping used to map the application characterization graph of the benchmarks aims to place traffic hotspots near each other. This explains why the gained throughput improvement is less significant than the ones attained with the synthetic traffics. However, mapping traffic hotspots in a close proximity is not always possible if bigger and more complex benchmarks are considered. Therefore, the global communication issue will continue to worry the designer and the proposed hybrid architecture is a very promising approach to tackle it.

Table 4.3: W-SWI PIR and throughput improvement over Mesh at the edge of network saturation where latency reach double the ZLL.

Traffic		Mesh	W-SWI	Improvements
Random	PIR	0.005	0.0073	46.0%
	throughput	0.062	0.089	45.0%
Transpose	PIR	0.0055	0.0081	47.3%
	throughput	0.068	0.098	43.8%
Butterfly	PIR	0.008	0.0096	20.0%
	throughput	0.098	0.116	19.1%
Shuffle	PIR	0.0075	0.009	20.0%
	throughput	0.092	0.109	19.2%
Hotspot1	PIR	0.0044	0.006	36.4%
	throughput	0.054	0.074	36.0%
Hotspot2	PIR	0.0045	0.0061	35.6%
	throughput	0.055	0.075	35.3%
Bittreversal	PIR	0.0062	0.0088	41.9%
	throughput	0.076	0.107	40.0%

Table 4.4: W-SWI average delay and throughput improvement over Mesh for a range of applications

Benchmark	No. of tasks	NoC size (Tile)	Improvement (%)	
			Average delay	Throughput
AMI49	49	7 × 7	45.5	19.3
AMI25	25	5 × 5	45.3	7.8
TELE	16	4 × 4	33.1	4.5

#### 4.4.2 Power Consumption

This section demonstrates results of the other projected major issue in global communication which is the increasing overall power consumption of interconnects system. The router's static and dynamic power is calculated using the Orion 2.0 [147, 122] model, and power dissipation for wire links is calculated for the x (3.6mm) and the y (5.2mm) directions. The modelled baseline router power has been calibrated to match the reported power measurement of the implemented NoC [5]. The transceiver (Tx/Rx) power consumption projection [18, 82, 101] is used (24mW per sub-channel). SWI power dissipation is also modulated based on the analytical model derived earlier in section 3.3.2. In addition, the DWA algorithm was designed using Verilog and then synthesized using Synopsys Design Compiler and mapped



onto the PDK 45nm technology library to calculate its dynamic power (6.97mW) and leakage power (21.95 $\mu$ W). Then all these values are used in Noxim to calculate the total power consumption of the overall network, as shown in Fig. 4.7.

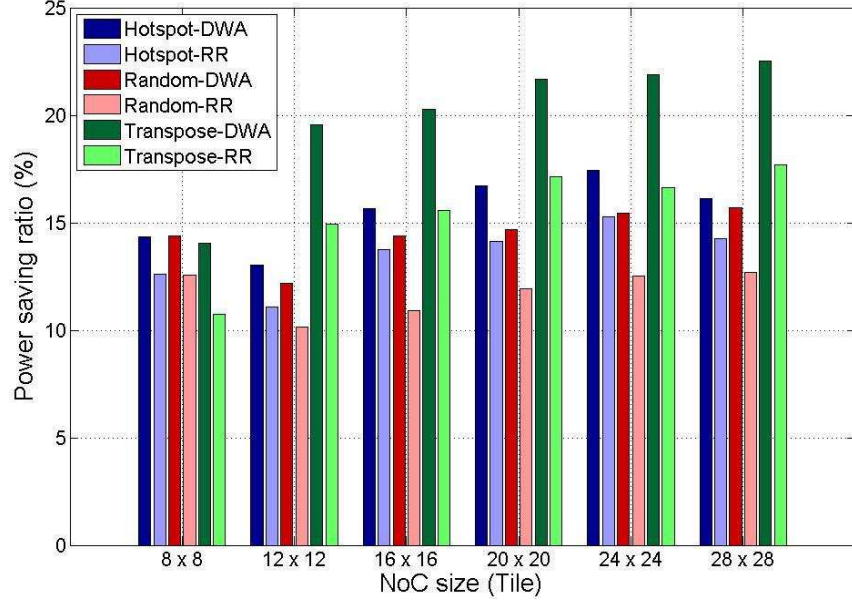


Figure 4.10: Comparison between W-SWI with DWA algorithm and with basic Round-robin (RR) Communication power saving ratio to the Mesh for different network size and traffic scenarios.

Fig.4.10 shows the power consumption saving ratio for the proposed hybrid architecture over the power consumption of the baseline architecture for different NoC sizes and synthetic traffic. It also compares between the proposed architecture when DWA algorithm used as a selection function with the same architecture when RR is used. It obviously demonstrates reductions in NoC power consumption by 12-23%, which is up to 6% enhancement over RR due to the DWA algorithm utilizing the proposed fabric more effectively. Moreover, it shows that as the NoC size increases for most cases the power saving ratio is increased due to the increasing distances between nodes, which increases the global communication problem. These findings prove that SWI has remarkable potential in mitigating the growing global communication issue.

#### 4.4.3 Area Estimation

It is essential to evaluate chip area overheads for the extra on-chip circuits required for our architecture. Firstly, for transceivers we assume that the active area calculated in [18, 82, 101] is the only part scaled down when we move to 45nm technology, while the passive part remains almost the same since they are proportional to the channels' operational frequency range. Therefore, the projected transmitter

Table 4.5: Area overhead consideration for the proposed Architecture comparing to related Architectures. Noticed that some component are needed only in the Master nodes

NoC component		Area per item ( $\text{mm}^2$ ) for 45nm technology		
Item	No. of items	Mesh	W-SWI (proposed)	RF-I with transmission line (TL)
Router	24	1.0853	1.5124	1.5124
Transmitter	6	-	0.1558	0.1558
Receiver	24	-	0.0083	0.0083
DWA	6	-	0.0044	0.0044
TL Link	1	-	-	0.7608
Wire Link	1	13.653	13.653	13.653
Total extra area over Mesh (all components $\times$ their numbers)		-	13.011	13.772
Total die area( $\text{mm}^2$ )		567.1		
NoC area (%)		7	9.01	9.15

area is ( $4870\mu\text{m}^2$ ) per sub-channel, while the projected receiver area is ( $260\mu\text{m}^2$ ) per sub-channel (the active area is proportional to the square of the scaling factor [148]). Secondly, the extra router port (buffer, crossbar and related circuits) area is calculated using the Orion 2.0 [147, 122] model to be ( $0.427\text{mm}^2$ ). The modelled baseline router area is a 6% less than the reported implemented router area [5], which is acceptable for the purpose of comparison evaluation in this study. In addition, the DWA algorithm was designed using Verilog and then synthesized using Synopsys Design Compiler and mapped onto the PDK 45nm technology library to calculate its area. On the other hand, the value of RF-I transmission line considered to be routed through the chip (NoC size  $6\times 4$ ) as (U) shape passing through all nodes to calculate its length[18, 82]. In addition, a transmission line with a pitch of  $12\mu\text{m}$  is considered in calculations of the link area overhead.

Table 4.5 shows area considerations for the baseline architecture, proposed SWI hybrid architecture and transmission line RF-I. Obviously, most of the extra area overhead for the W-SWI (2%) is due to the 6th router port and the increased size of the crossbar. Moreover, W-SWI offers a better area-performance trade-off compared RF-I transmission lines that offer the same connectivity [18], since transmission lines need to be implemented through the chip. Therefore, W-SWI actually saves the area that is needed by transmission lines in high metal layer.

#### 4.5 SUMMARY AND CONCLUSION

This chapter aims to explore potential scalable global communication features offered by the proposed hybrid wire-SWI architecture for on-chip interconnect, while maintaining a competitive area/performance trade-off. Zenneck surface wave low power dissipation and high signal propagation speed might allow significant advantage to address the global communication issues that are projected to scale up as the technology scales down. Therefore, SWI has been adopted for global communications, while enabling the regular wire to serve local and unicast communication. This has been accomplished considering these goals in two levels, the under-layer physical interconnects and the NoC multilayer topology. Moreover, simple DWA arbitration algorithms are developed that smartly and efficiently utilize the proposed architecture. Implementation considerations for this extra circuitry have been discussed. Results show significant improvements in terms of the average delay, throughput and power consumption without any topology optimization algorithm or task mapping. All that while having relatively small die area penalty comparing to the related work. As a result, the proposed architecture enabled by SWI has promising potential for future on-chip 1-to-1 global communication. The next chapter will investigate the potential of the proposed W-SWI architecture under one-to-many communication patterns.

## WIRE AND SURFACE-WAVE ARCHITECTURE WITH CENTRALIZED CONTROL FOR MULTICAST

---

### 5.1 INTRODUCTION AND MOTIVATION

As mentioned earlier, due to growing market demands, a growing burden on the NoC increasing significantly in terms of performance and power consumption scalability. This is especially true for CMPs, which have been introduced to provide near linear performance when complexity increases while maintaining lower power and frequency budgets [11]. This goal can be achieved by adopting the approach where many simple cores used instead of focusing on increasing single or multi-cores complexity [11]. Consequently, we are approaching a many-core era where the number of cores is expected to increase rapidly. Thus, a good utilization of such many-core architecture is becoming a crucial challenge.

The performance and power consumption of CMPs depend on both interconnect fabric and cache coherence protocols. These protocols rely heavily on an underlying communication fabric to provide one-to-many (1-to-M) communication. These protocols inject non-trivial percentages of 1-to-M packets [16, 15]. Moreover, some applications that are optimum for NoC-based CMPs, such as SNN modelling, inject high percentages of multicast traffic up to 100% into the NoC [116].

In the literature, NoCs conventionally treat 1-to-M traffic patterns as repeated unicast traffic, which is referred to as software multicast [3]. This basic handling of multicast by retransmitting the same data will increase power consumption and congestion. As a result, NoC performance is inversely proportional to the increase in the 1-to-M ratio, and even small ratios of multicast (1-to-M) or broadcast (1-to-all) will have severe effects on NoCs such as high latency and fast NoC saturation, see Fig. 5.1b. Therefore, NoCs designed for CMPs are expected to support multicast as well as unicast to meet applications demands.

Many studies have proposed wire-based NoC schemes that support 1-to-M communication. However, these studies struggle to match wire-latency and/or wire-energy [16, 15]. Nonetheless, these regular wire interconnects, which transmit the signal by charging/discharging the whole regular RC wire, find it difficult to maintain performance benefits in terms of latency and energy for even unicast global communication [7, 8]. Thus, these on-chip multicast schemes will not be sufficient in the near future. As a result, many researchers are looking for alternative communication fabrics, such as RF-based interconnects

[24, 18, 19, 22] and optical interconnects [27, 28]. However, although such interconnects seem promising, they might not be the ideal solution due to their complexity, incompatibility, power consumption and/or area overheads [7]. As mentioned earlier, The Zenneck surface wave [106, 105] is an emerging on-chip interconnect technology which is exploited in this chapter to mitigate 1-to-M communication issues. The major contributions of this chapter are as follows:

- To propose a fair and deadlock free arbitration and routing mechanisms for 1-to-M traffic that efficiently address multicast traffic and maximize W-SWI utilization. In particular, design the centralized arbitration techniques, that allow the concurrent utilization of many resources with relatively low circuit complexity and delay.
- To evaluate rigorously the W-SWI for both synthetic traffic and real application benchmarks. This proposed architecture is found to surpass the previous work by achieving improvements ( $\sim 22\times$ ) in average delay, ( $\sim 2 - 10\times$ ) in power consumption and attaining a reliable quality of service. Moreover, additional hardware cost of the W-SWI is found to be relatively insignificant.

#### 5.1.1 Motivation

In literature, NoC conventionally treat 1-to-M traffic patterns as repeated unicast traffic (software multicast) [3]. This basic handling will have a dramatic effect on the NoC, for the following reasons:

1. 1-to-M increases congestion on the source node of this traffic (router, network interface and links) and thus creates bottlenecks.
2. It causes a poor QoS due to the queueing of repeated unicast packets on the same communication fabric. Thus, it is difficult to provide guaranteed service.
3. Power consumption is increased due to retransmitting the same data, but to different destinations.

As a result, even small percentage of 1-to-M traffic will have severe affects on NoC performance and cost , as shown in Fig.5.1b.

Cache coherence protocols depend on a range of 1-to-M communication patterns such as multicasting invalidation requests (directory-based protocols) and broadcasting ordering tokens (broadcast-based protocols) [15, 16]. Cache coherence broadcast-based protocols (such as token coherence) offer less hardware overheads than directory storage that scales with number of cores, and low latency unlike other cache coherence protocols [16, 149]. However, in these protocols the ratio of broadcasting (a special case of 1-to-M) to the total PIR is considered

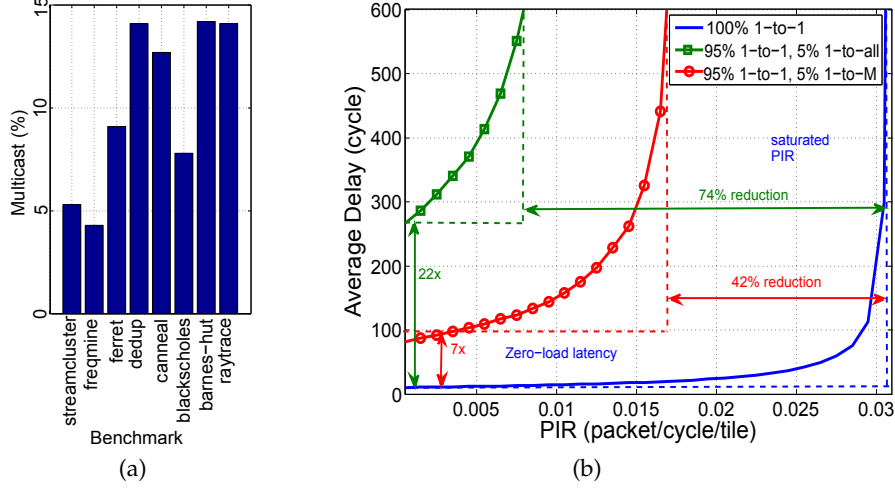


Figure 5.1: (a) The non-trivial 1-to-M traffic percentage according to the simulation of a range of CMP benchmark applications (from PARSEC and SPLASH2) with MESI cache coherence protocol; (b) our  $6 \times 6$  regular mesh NoC simulations with random traffic plus random traffic with a small percentage of multicast or broadcast (5%). The introduction of multicast or broadcast leads to severe deterioration in performance in terms of latency and saturation PIR.

to be very high (5% to 52.4% [16, 15]). For instance, Fig. 5.1a shows multicast ratios for a set of standard benchmark applications from PARSEC [150] and SPLASH2 [151]. All these benchmark applications were running with the MESI cache coherence protocol.

This could be catastrophic for global coherence unless the interconnect fabric supports the 1-to-M communication. Therefore, the trend in cache coherence protocols design is to mitigate (because they are unable to eliminate) 1-to-M communication. For example, the multicast injection ratio ranges from 3.1% to 12.4% [15]. As a result, broadcast-based protocols, or any other promising solutions that require high ratio of 1-to-M and/or large multicasting destination groups, are avoided. This study attempts to eliminate these constraints and improve performance by proposing an interconnect architecture using the emerging surface wave technology.

Moreover, the NoC-based CMP has been found to be naturally suited for applications such as SNN modelling since SNN requires high processing and communication parallelism [152]. These networks have potential in many applications, such as mimicking mammalian brains to solve complex intelligent tasks and medical applications such as to replace damaged brain cells. The key aspect of this type of SNN is the real-time performance of the communication architecture because they depend on to match the behaviour of biological neurons. Moreover, a scalable architecture with low power and area budgets is also a vital feature for future spiking-neural-networks (SNNs). These requirements

represent a challenge even with full custom designs [153], since the multicast rate in such networks may reach 100% with high graph density.

Most previous studies have suggested tree-based, see Section 5.2.1, area/power hungry look-up tables for SNN [116, 152]. Some promising architectures leverage the communication nature of SNNs by proposing hierarchical NoCs that support local and global interconnects in different interconnect layers [153]. However, these are still limited by global wire/router fabric latency. Thus, this study proposes an upper global interconnect layer that overcomes these issues, which other custom (or general-purpose) NoC-based CMP architectures and schemes struggle to cope with.

## 5.2 RELATED WORK IN CURRENT AND EMERGING INTERCONNECTS FOR MULTICAST

In this section a trend in wire-based multicast routing and architectures will be outlined and reviewed, which will raise the need for new emerging interconnects. Thus, the fanout feature of promising emerging interconnects will be discussed. In addition, examples of state-of-the-arts interconnects architectures that support multicast and enabled by these emerging interconnects are going to be reviewed.

### 5.2.1 *Wire-based Multicast Routing and Architectures*

Previous studies in 1-to-M on-chip communication have tried to solve the challenges of basic software-multicast routing by altering the interconnect fabric hardware and/or introducing multicast routing algorithms. These studies can be divided into path-based and tree-based schemes. Firstly, in a path-based scheme, a 1-to-M packet is submitted to members of a multicast group sequentially following one path and dropping a copy at each destination node [154], as shown in Fig. 5.2a. This method provides a non-trivial saving in energy and less load on the interconnect fabric, since it propagates one packet, compared to a software multicast. However, it offers varying and very significant delivery latency, especially in broadcast scenarios. Moreover, careful consideration might be needed to avoid creating a deadlock scenario due to the packet worm non-minimal path.

Secondly, in tree-based schemes, 1-to-M traffic is propagated in a tree route embedded in the network that branches only at the disjoint points in the path [16, 15], as shown in Fig. 5.2b. This alleviates traffic load near the source and reduces energy consumption by allowing the 1-to-M traffic to be duplicated only when it has to fork. However, there are few issues facing tree-based routing. The first issue with tree-based routing is the multicast dependency that occurs sometimes between two or more 1-to-M wormhole traffic requesting simultaneously the

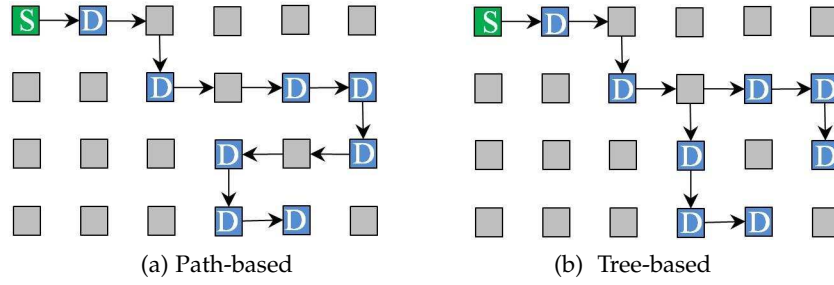


Figure 5.2: Demonstration of different 1-to-M routing schemes: (a) path-based delivers packets sequentially in a worm-like route 5.2a, (b) Tree-based delivers the packets in a tree-like route [16, 15].

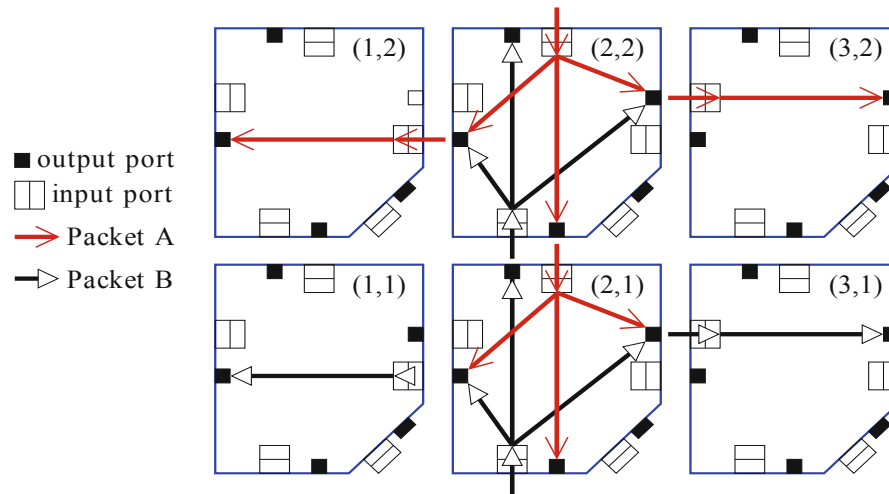


Figure 5.3: Illustration example of multicast dependency that causes deadlock in tree-based multicast routing [156].

same output ports, as shown in Fig. 5.3. Therefore, these tree-based routing algorithms are limited to packet switching [15]. Nonetheless, a novel tree-based multicast routing for wormhole flow-control has been proposed [155, 156]. Such deadlock-free multicast routing algorithms avoid the multicast dependency by dynamically, locally, and virtually allocating each port based on a flow tag. However, this would require non-trivial complex control circuits and reservation tables to interleave flows flit by flit. In addition, it is limited to routers where their micro-architectures consist of distributed routing and arbitration units over each output port.

On the other hand, although tree-based routing reduces congestion near the traffic source, the second issue with tree-based routing is duplicating the packets in intermediate nodes. This way a packet is forward to disjoint destination(s), which adds non-trivial power



consumption. To handle this issue and balance between the path-based and tree-based routing, a set of mixed tree-path-based routing algorithms has been proposed [157, 158, 154]. These methods are either deterministic [154] or partially adaptive [157, 158] routing algorithms. In this type of multicast routing algorithms, the destinations are divided into subsets based on their location. Then the packets are duplicated only in the source node (the tree root) so that each copy is forwarded in a path-based style to each sub-group of multicast destinations. Nonetheless, the path-based avoids deadlock by following a Hamiltonian path to visit each destination. However, the packets are still duplicated up to four times and routed in some cases via non-minimal path. Therefore, it still suffers from performance and power consumption issues. Moreover, this multicast handling does not fully address the challenging issues of the regular metal-based NoCs, since it mitigates but does not eliminate the wire issues, excessive extra load, and hotspots that reflect significantly on network performance and power consumption [159].

### 5.2.2 *Optical interconnects*

The interconnect infrastructure should support 1-to-M communication to cope with future many-core requirements, as mentioned earlier. Therefore, although the optical-interconnects does not offer a natural fanout feature, many studies have proposed optical interconnect architectures for multicast. These studies either suggest free-space or waveguided optical interconnects. The free-space optical interconnect directs the signal using chip surface devices such as micro-lenses, micro-mirrors, diffractive optical elements (DOEs), laser sources and photo-detectors (PDs) [75, 28]. However, only a few studies have investigated the option of optical free-space for clock distribution, but not for data multicast [28].

On the other hand, waveguided optical interconnects have been thoroughly investigated and many state-of-the-art architectures have been proposed. These vary in topology and the on-chip devices that support them. For example, the tree-topology requires splitters and combiners to fork and join the optical signals [160], as shown in Fig. 5.4a. Another example is a bus-based topology that utilizes wavelength-division-multiplexing (WDM) and then uses a bank of microring modulators, which can be configured to listen to a selected channel [29], as shown in Fig. 5.4b. However, all these architectures have limited fanout capability because the optical signal would decay significantly after each forking or partial drop of the signal to a receiver node [29, 160]. The number of nodes that can receive the signal depend on the signal power budget, which is considered to be relatively high.

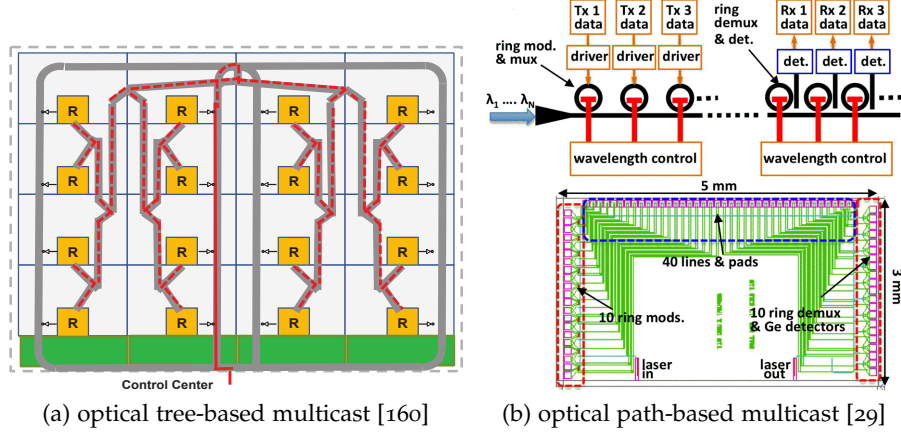


Figure 5.4: Optical-based interconnect architecture examples that support on-chip multicast.

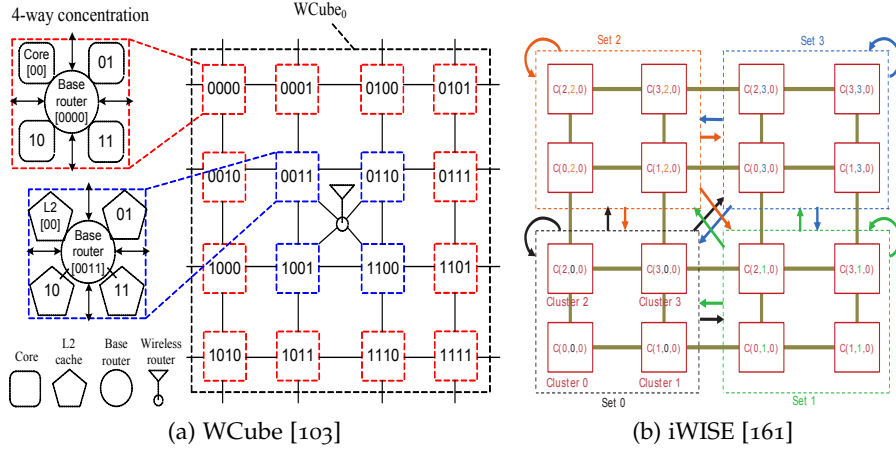


Figure 5.5: Examples of multicast clustering in WiNoC-based interconnects architectures.

### 5.2.3 Wireless Interconnects

In contrast to optical interconnects, the WiNoCs have natural scalable fanout capability, which makes them preferable for multicast-enabled interconnect architectures. As a result, many studies have suggested the WiNoC for CMPs with multicast requirements [23, 103, 161]. However, the WiNoC fanout capability depends on the antenna radiation pattern and coverage distance, which are up to 23mm [22]. This is due to high power dissipation of the RF signals when they propagate in free space, which leads to a low coverage distance to power ratio. Therefore, the transceiver power amplifier and the antenna design should take into consideration the required distance and the directions of the destinations. For instance, some studies propose run-time tunable transmitting power based on the required destination [93].

In terms of connectivity, most researchers have suggested a virtually 1-to-all connectivity for each RF-transmitter node in the wireless interconnect layer [81, 23]. This does not mean that all the nodes are able to communicate with all other nodes simultaneously. However, these RF-transmitter nodes are competing over the shared media. Therefore, contention issues are a main challenge for such architectures and especially when facing the dramatic scalability of NoC size [23]. The other type of multicast architecture depend on NoC clustering, where each cluster, either statically or dynamically, would be listening for a specific carrier frequency [161, 103], as shown in Fig. 5.5. Thus, this clustering should mitigate contention, but would increase reconfigurability and routing complexity. Moreover, this approach limits the crucial fanout capability to a subset of the nodes.

#### 5.2.4 *Transmission Lines*

Although the RF-I has a low ratio of power dissipation to signal propagation distance, RF-I-based multicast architectures face several challenges. For instance, RF-I tree-topology forking requires stubs, which means an impedance discontinuity. Therefore, a careful matching circuit design is required at the end of each stub [1]. This would increase design complexity especially if the stubs lengths and distribution of forking points are non-uniform.

Therefore, to avoid using a tree of TLs, many designs have proposed a worm or cycle layout of this thick wires to pass through all the nodes, as shown in Fig. 5.6. This layout involves further challenges such as adding nontrivial area overheads, signal decay and signal latency. Firstly, in terms of area overhead, the signal distribution in RF-I is limited to the nodes that transmission lines pass by them. As a result, the worm or cycle layout of these thick wires should go through almost every tile in the chip [82, 18, 19]. This might add nontrivial area overheads and on-chip routing issues because the pitch of TLs (width and spacing) is relatively large. Secondly, this layout might mitigate but not eliminate the impedance discontinuity unless careful matching circuit is introduced. Therefore, multicasting the signal to many destinations is not scalable because, with each drop point, the signal decay, latency and signal reflections are increased unless careful matching circuits are designed [21, 1].

#### 5.2.5 *SWI Fanout Feature*

Unlike all the already mentioned emerging interconnects, SWI interconnect offers natural efficient fanout features. For instance, the E-field decay rate in SW from the source horizontally along the boundary should be around  $(1/\sqrt{d})$ , as mentioned in chapter 3. On the other hand, vertically, the decay is exponential away from the boundary.

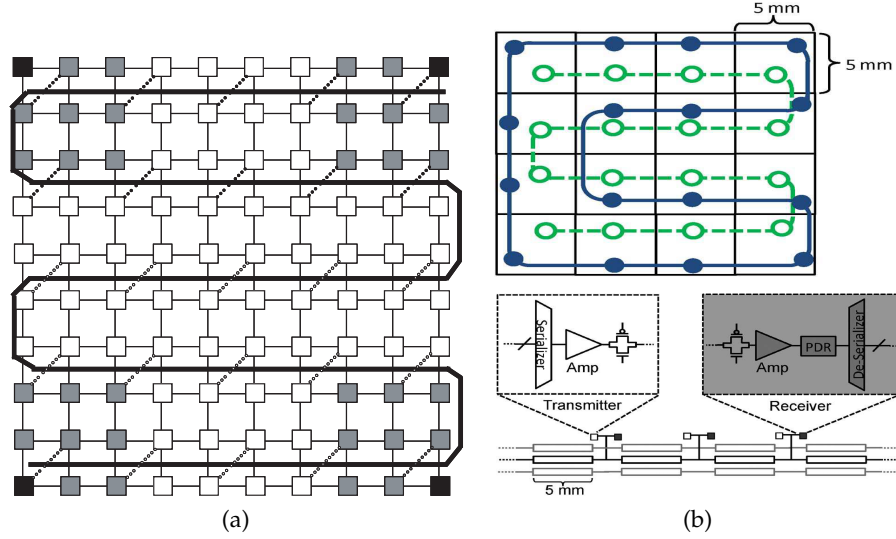


Figure 5.6: Examples of some RF-I multicast architectures [18, 19].

This allows less power dissipation for far larger coverage areas than the regular WiNoC since they for the same SNR propagate the signal up to 10cm [35], 23mm [22, 94], respectively. This is due to the fact that RF wireless signals are dissipated via antennas and free space. However, both WiNoC and the SWI signals transmit in all directions (over the surface for the SWI) at a speed close to the speed of light if we assume the WiNoC antenna radiation pattern is circular  $360^\circ$ . Thus, SWI can fanout the signal across the chip in one clock cycle with competitive levels of power consumption and circuit complexity compared to other emerging interconnects [35].

However, like wireless, there are two cases in SWI multi-receiver of the signal: The first when they are on different angle from the source, then the transducer, similar to the antenna in wireless, would receive the electromagnetic signal without effecting each other. The second case is when they are on the same or close angle of the source. Then each transducer would absorb part of the signal power. Thus, fan-out is limited by the transmission power if the receivers within the same direction. However, the impedances will remain the same since every transducer that has been designed to match the impedance of the surface will absorb/receive part of the electrical-field and the magnetic-field (the ratio between them remain the same) and waveguide surface characteristics that define the impedance remain the same. As a result, SWI does not suffer from the same discontinuity like the transmission-lines with stubs. Therefore, SWI may be the favoured supplementary interconnect for future CMPs that depend highly on 1-to-M.

### 5.3 W-SWI MULTICAST ROUTING SCHEME

Multicast traffic benefits from the characteristics of the SWI, but a smart routing technique is required to direct and deliver multicast traffic to its final destinations and to maximise the benefits at minimum cost. In this thesis, the 1-to-M routing for the proposed architecture is an improved tree-based scheme where the embedded tree path forks at one point; specifically, the nearest master. Therefore, the maximum degree of routing graph is up to  $(N - 1)$ ; thus, the need for the SW fanout feature. The nearest master then delivers concurrently the flit to all the addressed leaves (slaves) in one hop via SWI as illustrated in Fig. 5.7. This is accomplished by adding the following rules to the routing rules (*Rule 1* and *Rule 2*) presented in Section 4.3.2:

- *Rule 3*: Direct all the 1-to-M traffic to the nearest master using any deadlock-free routing algorithm.
- *Rule 4*: If the master is part of the multicast group, the packet cannot request or allocate the router local port until it has used and released the output port linked to SWI.

These rules provide higher efficiency in handling 1-to-M traffic and avoiding deadlock, for the following reasons. Firstly, by *Rule 3*, each node simply needs to direct the 1-to-M traffic to the nearest master using any simple deadlock-free routing algorithm. For instance, as mentioned in section 4.3.2, a simple partially adaptive algorithm such as the odd-even could be adopted since it offers more path diversity than other turn model routing schemes [55]. Hence, there is no need for extra circuitry or complicated algorithms to build the multicast tree path and to determine the forking points. Moreover, this routing scheme allows the tree to have only one forking point (the nearest master), and due to the fanout feature of the SWI, packets will be replicated only at the destination routers. This will reduce power consumption by eliminating the need for duplicated traffic to travel through costly (power hungry and already loaded) intermediate wires and routers.

Secondly, by adding *Rule 4*, we would insure that the multicast dependency could be only created in the SWI layer. This multicast dependency is a major issue in multicast routing because it leads to deadlock scenario. The multicast dependency in SWI layer issue and its handling scheme will be presented in Section 5.4.

In order to direct the packet to the multicast group members, each multicast packet header must have multicast-address-bits (MAB). This header field is a bit vector where each bit represents a node, and it is set if the node is a member in the multicasting group. This header field might also be used along with the source identity (ID) to uniquely identify a 1-to-M request, as will be described in the following section. This multiaddress encoding technique is far more efficient

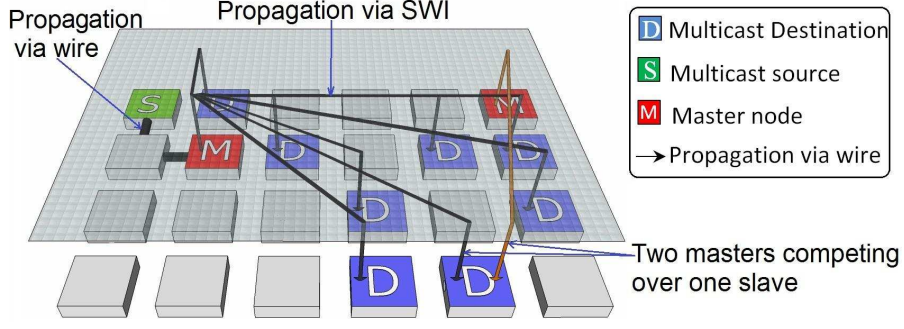


Figure 5.7: W-SWI improved tree-based with low latency packet delivery where branching possible only at the SWI master nodes .

than all-destination encoding and simpler than region-encoding [3]. Moreover, another field that is added to the packet header is the multicast/unicast bit, which allows the routers to distinguish between a multicast and unicast packets. This way unicast packets header does not need to have MAB field. As a result, it reduces the header size overhead. After the flit has been successfully delivered to slaves via the SWI, it will wait until it has been drained locally by the master (if the master is part of the multicast group) in the next clock cycle before being ejected from the router buffer. this considered the only forking point happening in the network mesh layer.

#### 5.4 PROPOSED ARBITRATION AND ALLOCATION SCHEMES

This section presents the arbitration and allocation challenges in the SWI layer of the proposed NoC hybrid architecture. Design rational and discussion are then offered for the centralized approach to address these challenges.

##### 5.4.1 Contention Challenges

Arbitration and allocation is crucial to avoid channel starvation and deadlock in the SW layer. The SWI can be represented logically as a bus topology with multiple master nodes with Tx/Rx capability and multiple slave nodes with Rx capability. Each master has its own dedicated logical bus (frequency channels). However, each slave can receive from one master at a time, which creates competition between masters. This competition escalates as the size of multicast group members (destinations) and the number of masters increase. As a result of this competition, two scenarios might develop: channel starvation and a multicast dependency deadlock. The first issue results when master(s) win the allocation of slaves repeatedly while other master nodes are waiting.

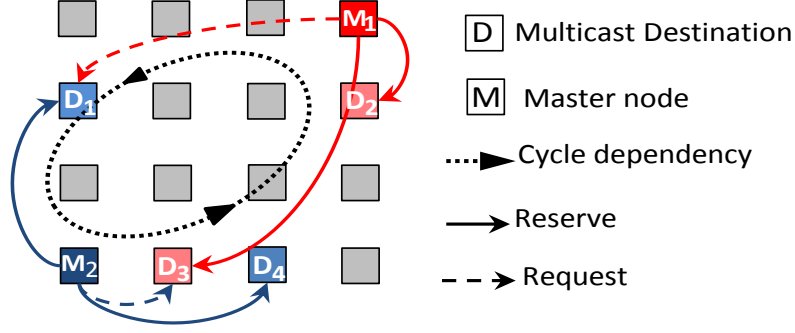


Figure 5.8: Demonstration of the deadlock problem created by the multicast dependency.

The second scenario could occur if one master ( $M_1$ ) allocates only a subset of the multicast group (slaves) because the rest of the requested slaves have already been allocated. In the next arbitration,  $M_1$  might lose the arbitration of the remaining subgroup to another master ( $M_2$ ). Consequently, if  $M_2$  is requesting a subset of the slaves allocated by  $M_1$ , a cycle dependency will occur. For example, Fig. 5.8 demonstrates the deadlock problem resulting from multicast dependencies between two masters  $M_1$  and  $M_2$  requesting slaves  $\{D_1, D_2, D_3\}$  and  $\{D_1, D_3, D_4\}$ , respectively. In this example,  $M_1$  manages to allocate  $\{D_2, D_3\}$  and now is waiting for  $D_1$ , while  $M_2$  allocates  $\{D_1, D_4\}$  and now is waiting for  $D_3$ . This scenario will cause a deadlock situation, since each master will not release its allocated slaves unless it delivers the rest of the packet flits.

Moreover, the use of virtual-channel for surface-wave-interconnect (VSWI) offers the best utilization of SWI and enables slaves to listen virtually to more than one master. This way, if a master is waiting for a message to be delivered to the rest of slaves or this message needs to be drained locally by the same master, idle reserved slaves will not be prevented from accepting traffic from another master. This complicates the allocation problem in finding a legal match between masters  $\times$  virtual-channels for surface-wave-interconnect (VSWIs)  $\times$  slaves (three dimensions) so that no two masters are allocated the same VSWI for the same slaves simultaneously. Therefore, to offer fair deadlock-free arbitration while efficiently utilising the W-SWI, a centralized approach is proposed in the this chapter.

#### 5.4.2 Centralized Arbitration

This section describes the design of the global-multiresources-arbiter (GMA) and its rationale in addressing the contention problems mentioned above in a centralized approach. The resulting hybrid architecture with GMA will be referred to as W-SWI centralized (W-SWI-C). To avoid the multicast cycle dependency scenario mentioned earlier,

slaves could be allocated as a group. This can be achieved in the arbitration request masking stage by using the MAB-Check unit, which shown in Fig. 5.9. This unit will not validate a request from any master unless all the slaves that have been requested are free by comparing the request with the content of a GMA reservation table. In this way, the arbitration problem is also minimised from the three-dimensional matching masters  $\times$  VSWIs  $\times$  slaves to the two-dimensional matching masters  $\times$  VSWIs.

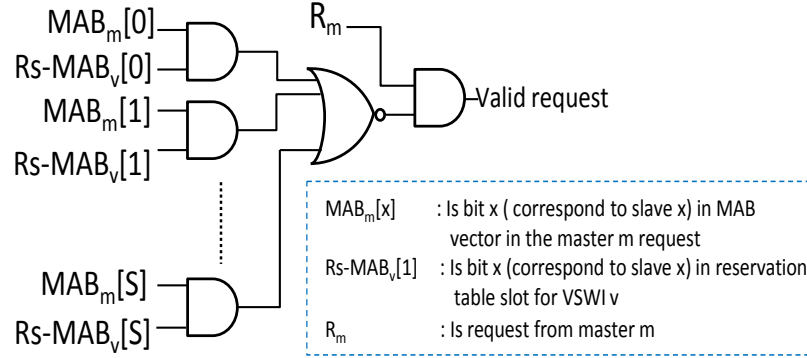


Figure 5.9: The proposed MAB-check unit to mask the requests unless all the requested multicast group are free.

Arbitration in systems such as CMPs requires a high parallelism feature in order to minimize decision latency. Moreover, a legal matching where an output is assigned to one input and vice versa is crucial in multi-resource arbitration. These two features can be achieved between the vector of inputs, which represents the masters, and a vector of outputs, which represents the VSWI, by using two sets of arbitration: one for the input and one for the output. However, this is likely to lead to a poor legal matching or minimum matching, where less than optimal possible resources have been allocated. Optimum legal matching (maximum matching) can be achieved by adopting a lonely output allocator (LOA) that introduces one more stage before input arbitration [51]. This extra stage counts the number of valid requests for each output in order to detect the level of competition over each output. Then, the less popular output (or least-requested) will be given higher priority in the next arbitration stage. This should minimize conflict and produce maximum matching whenever possible.

Fig. 5.10 and Algorithm 5.1 show the structure and the operation of the proposed GMA which achieves the best legal match in two cycles



---

**Algorithm 5.1** procedure of centralized arbitration and allocation using GMA.

---

```

.....Stage 1 -
1: for all  $i \in \{0, \dots, N_m - 1\}$  do
2:   for all  $j \in \{0, \dots, N_v - 1\}$  do
3:     if  $R_i$  then
4:        $E_{i,j} = !((MAB_i[0] \vee Rs-MAB_j[0]) \wedge (MAB_i[1] \vee Rs-MAB_j[1]) \dots \wedge$ 
          $(MAB_i[N-1] \vee Rs-MAB_j[N-1]))$ ;
5:     end if
6:   end for
7: end for

.....Stage 2 -
8: for all  $j \in \{0, \dots, N_v - 1\}$  do
9:    $C_j = \sum_{i=0 \rightarrow N_m-1} E_{i,j}$ ;
10: end for
11:  $P = \min(C_0, C_1, \dots, C_{N_v-1})$ ;

.....Stage 3 -
12: for all  $i \in \{0, \dots, N_m - 1\}$  do
13:    $X_{i,0} = \text{VariableArbiter}((E_{i,0}, \dots, E_{i,N_v}), P)$ ;
14: end for

.....Stage 4 -
15: for all  $j \in \{0, \dots, N_v - 1\}$  do
16:    $\text{minTimeS} = \text{CurrentTime}$ ;
17:   for all  $i \in \{0, \dots, N_m - 1\}$  do
18:     if  $(X_{i,j} \text{ and } \text{TimeS}_i \leq \text{minTimeS})$  then
19:        $\text{winner} = i$ ;
20:        $\text{minTimeS} = \text{TimeS}_i$ ;
21:     end if
22:   end for
23:    $W_{\text{winner},j} = 1$ 
24: end for

.....Stage 5 -
25: for all  $j \in \{0, \dots, N_v - 1\}$  do
26:    $\text{RT}(W_{\text{winner},j}, \text{MAB}_{\text{winner}})$ ;
27: end for

.....Stage 6 option 1 -
28: for every Allocation Cycle do
29:    $G = \text{RR}(\text{RT})$ ;
30: end for
31: if any  $H_{ID} = \text{Release}$  then
32:    $\text{RTrelease}(\text{MAB})$ ;
33: end if

.....Stage 6 option 2 -
34: if any  $H_{ID} = \text{Release}$  then
35:    $\text{RTrelease}(\text{MAB})$ ;
36:    $G = \text{RR}(\text{RT})$ ;
37: end if

```

---

(given that the requested resources are free) and with remarkably low circuit complexity. When the request is received from masters via SWI, they are demodulated and the request data are extracted, such as the requested destination(s) (MAB), a time-stamp and a source-ID. The first stage is a request masking in order to check if the master's request is possible for any of the VSWI by comparing its MAB with already reserved resources in the reservation table using the MAB-check unit, as mentioned earlier. The next three stages (2-4) represent the LOA which achieves the maximum matching of master-to-VSWI by minimising conflict between master requests over VSWIs. This is accomplished in stage 2 by counting the valid requests for each VSWI, and then generating the priority signal that will prioritize a VSWI that is subject to less competition. Afterwards, in stage 3, each request will elect one of the VSWIs to compete over it. The elected VSWI should have less competition over it and the slaves requested by this master are currently free. In stage 4, the final stage of LOA, requests are competing for each VSWI, where the winner will be out of comparator tree arbiters is the oldest.

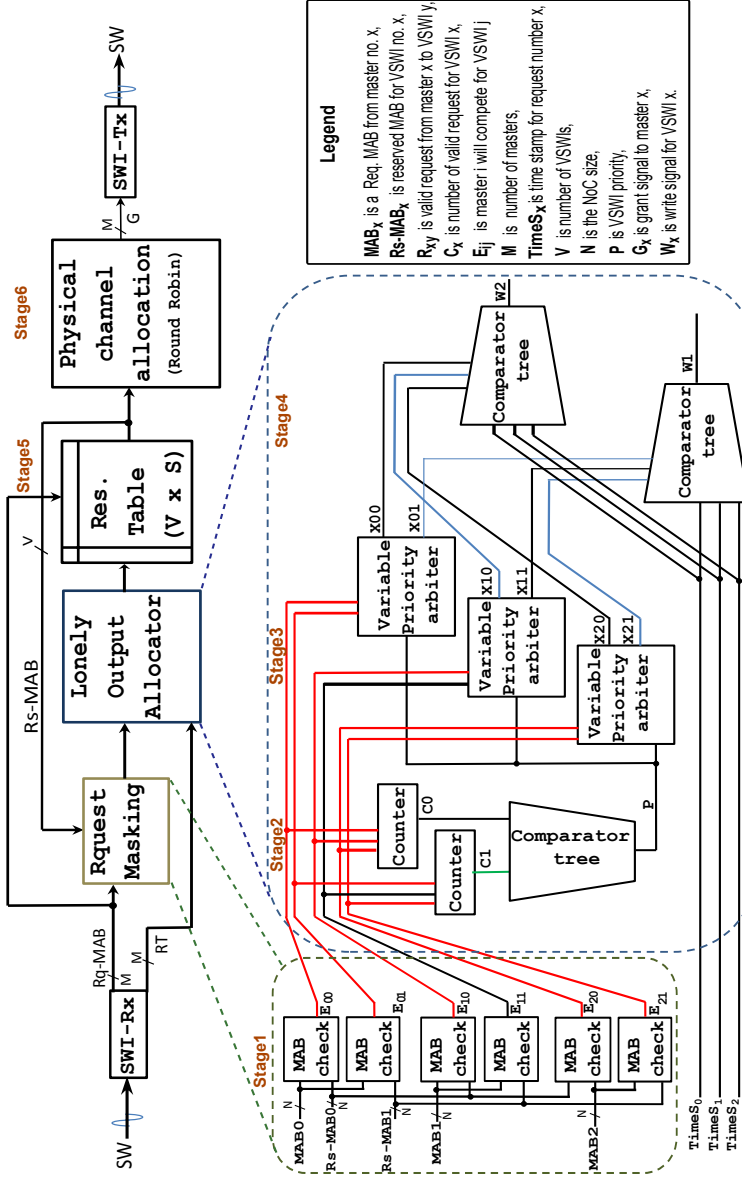


Figure 5.10: Design of the proposed global multi-resource arbiter (GMA) for SWI channels: stage (1) request masking; stages (2-4) achieve legal match with lonely output allocator [51]; stage (5) generates the grant signals for a fixed period. The figure also shows an example of GMA stages 1-4 with four masters and two VCs. Master ( $M_4$ ) related logic is not drawn, for simplicity, but it is currently allocating some of the slaves requested by  $M_2$ .

The LOA is followed by stage 5, where the winning request from earlier stages will be stored in a reservation table. The size of this table is proportional to the NoC size, the number of master nodes (SWI channels), and the number of VSWIs. The final stage 6 represents physical channel allocation, which switches among subgroups of slaves reserved in previous stages under the same VSWI in order to generate a grant signals for a slaves subgroup in a limited number of clock cycles. This stage utilizes the VSWI to provide higher performance by allowing non-conflicting masters to transmit at the same time using their own channel frequencies. This stage consists mainly of a simple arbiter such as a RR arbiter. The duration of the allocation can be designed so that the allocation period is:

1. One cycle.
2. Fixed period, which would need a frequency divider.
3. As long as the request is asserted, which would need a hold release mechanism [51], in short, Hold.

Tuning between these options mainly depends on traffic pattern and system-level evaluation, as discussed in Section 5.5.1.

The final step is where the output is stored in an allocation register, and will be transmitted as a grant signal through an SWI-specific frequency control channel. The time cost of the arbitration in case of winning the arbitration is two clock cycles. Otherwise, the delay would equal to the arbitration cost plus a blocking period ( $T_b$ ), which could be up to  $\text{Packet}_{\text{size}} \times N_v$ , where  $N_v$  is the number of VCSWs if there is no congestion in the slaves.

To illustrate the GMA functionality, Fig. 5.10 also shows an example of the first four stages of the GMA that serves four masters with two VSWIs. However, the logic related to the forth master ( $M_3$ ) is not shown, for simplicity, because it is currently merely reserving  $S_3$  via VSWI1. Masters  $M_0, M_1$ , and  $M_2$  have requested the slaves  $\{S_1, S_2\}$ ,  $\{S_2, S_3\}$ , and  $\{S_1, S_2\}$  respectively. Since  $S_3$  is already reserved via VSWI1,  $E11$  is the only signal deactivated (not color red) by the MAB-check unit. In stage 2, therefore, the output priority (P) will be for the competition over the less-requested VSWI, which is VSWI1. Then, through the priority arbitration in stage 3,  $M1$  and  $M3$  will compete over VSWI1 while  $M2$  will compete over VSWI0 alone, which are highlighted in blue. The winner of stage 4 (W1) will be the master request with the lowest value of the time-stamp (TimeS).

#### 5.4.3 Communication Protocol

The communication protocol among masters, slaves and the GMA taking place at the SWI level is shown in Fig. 5.11. The master interface sends a request on the same master data frequency channel (channel

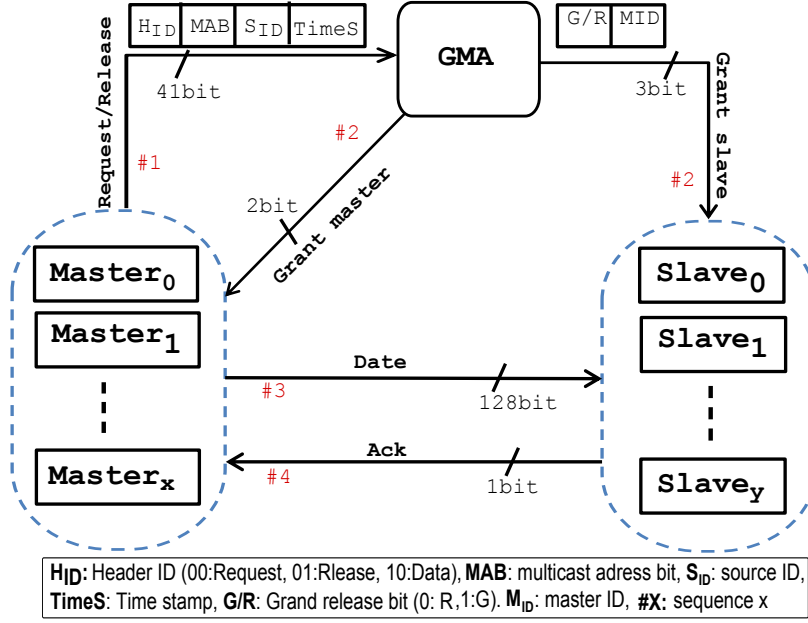


Figure 5.11: Demonstration control signals of SWI communication protocol exchanged between master, slaves and the GMA.

establishing phase). This request is identified by the header ID ( $H_{ID}$ ), which distinguishes received flit to be either request, release, or data flits. In addition, the request packet consists of the data required for the arbitration process, such as the MAB, TimeS and Source ID ( $S_{ID}$ ).

When the arbiter grants the request, it will generate two types of signals. The first type is the master grant signal, which grants their requests for the next cycle. The second grant signals are to be sent to the slaves. These signals consist of two parts: the grant/release bit (G/R) and the requested master number ( $M_{ID}$ ) that will inform the slave which channel it should listen to. After these signals are received, the data handshake phase starts. The master transmitter sends a data flit and waits for all the slaves in its multicast group acknowledge signals. In order to utilize the limited available frequency bandwidth, the master request signal is transmitted via the same dedicated master data channel frequency. Thus, the rest of the control signals require a bandwidth of 104 bits (for 4 masters and 24 slaves). The frequency bandwidth available for 45nm technology should be sufficient for the bandwidth required by the four data channels and the extra control, see Chapter 3.

## 5.5 SYSTEM LEVEL EVALUATION AND DISCUSSION

This section presents results obtained from our cycle-accurate NoC simulator which was built by modifying the existing Noxim simulator [121] for the W-SWI, VCT and the baseline architecture of regular mesh; in short Mesh. The Mesh in this chapter refers to a wire-based mesh

NoC that manages the 1-to-M traffic as a software multicast. In this thesis, the Intel SCC [5] is adopted as the baseline architecture as mentioned earlier. This chip is designed for performance critical CMPs, which makes it optimal for the purpose of this study.

Packet sizes of 1, 4 and 12 flits were chosen as an example to demonstrate the behaviour of the proposed architecture under packet-switching, virtual-cut-through-switching, and wormhole-switching flow control, respectively. The number of master nodes is based on the available frequency range for 45nm, which was estimated to be four channels (plus the frequencies specified for control signals). However, this frequency range is scaling with technology [18]. As a result, the number of master nodes was increased when simulating larger NoCs, assuming that the technology will have been scaled too. In addition, a VSWI number of four was chosen, which realizes a better performance/cost trade-off. In addition, in the evaluation in this section the VC was chosen to be equal to the VSWI for simplicity of router architecture.

The simulation was conducted with synthetic traffics, which are: (1) Random, where packets are transmitted randomly with uniform probability to other nodes; (2) Hotspot, which is the same as the random but with specific nodes called hotspots, four in this case, with higher probability of traffic dispatched to them; (3) Transpose, where a node sends a packet to other node that has its address transposed. These synthetic traffic adjusted to inject a specific percentage of broadcast (1-to-all) or multicast (1-to-M). The source nodes of this multicast traffic are selected randomly during the simulation, while the rest of the traffic consists of normal unicast packets according to the named synthetic traffic. In addition, in the case of multicast, the destinations of these packets are also selected randomly.

#### 5.5.1 Performance Improvements

This section presents a performance evaluation of the proposed architecture under synthetic traffic. Fig. 5.12a shows much less average delay for the W-SWI-C over the Mesh with Random traffic consisting of 10% broadcast and 90% unicast. Even for ZLL, the average delay improvements is  $\sim 22\times$ . Similar improvements are obtained with Transpose ( $\sim 24\times$ ) and Hotspot traffic ( $\sim 21.8\times$ ). These significant improvements are due to the software multicast in Mesh that replicates the multicast traffic to all destinations. Thus, it will increase the load and hotspots on the NoC. In contrast, In W-SWI-C, packets are replicated at the multicast destination routers and through short-cut links that avoid costly intermediate routers and wire links. In addition, when multicast traffic is used, the resulting improvements were up to  $12\times$  over Mesh. Obviously, this is due to the lower load and hotspots caused by multicast rather than broadcast.

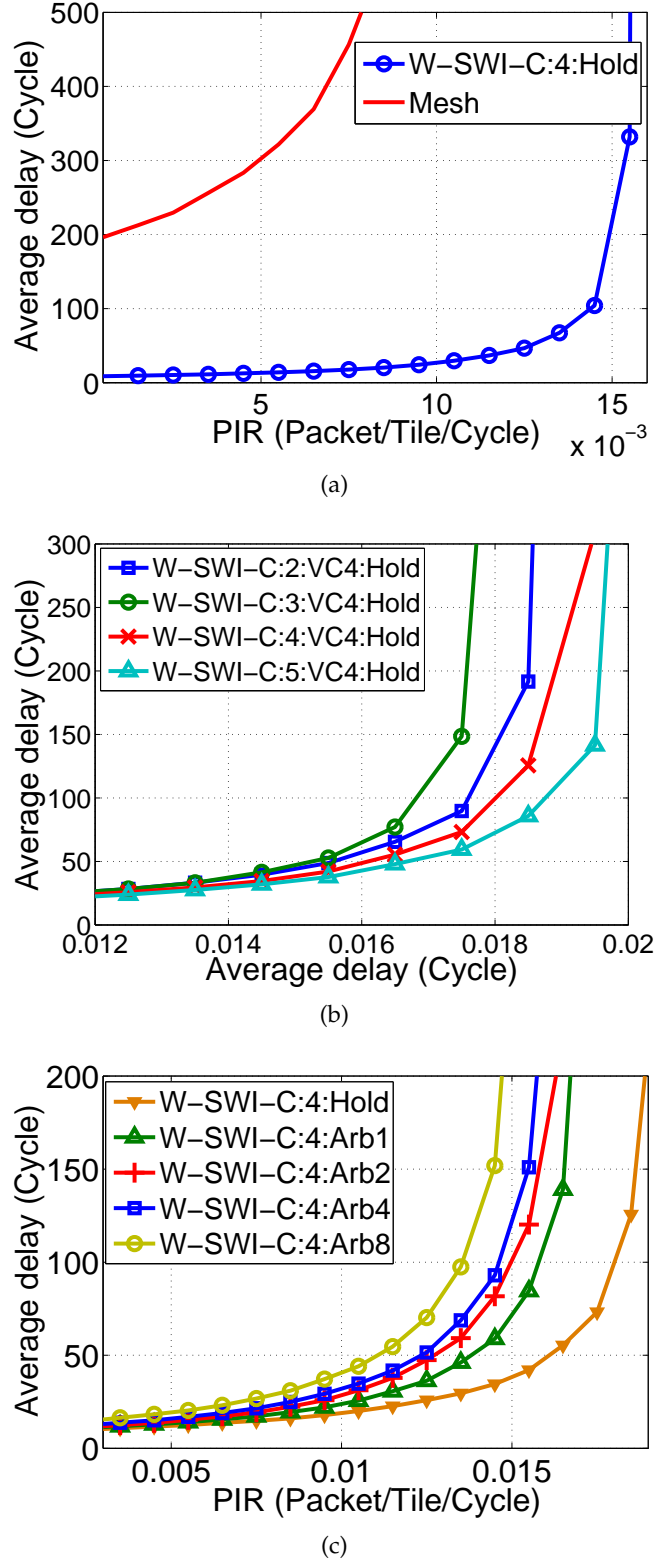


Figure 5.12: Average delay results of  $6 \times 4$  NoC with the following: (a) comparison of Mesh and W-SWI-C under random traffic with 10% broadcast; (b) W-SWI-C with different allocation techniques; (c) W-SWI-C with different SWI master number. Note that W-SWI: $N_m$ : $VCN_v$ :ArbP refers to W-SWI with  $N_m$  number of masters,  $N_v$  number of VC and P number of grant cycles. Hold is where GMA grants a master until all its current data flow is transmitted.

On the other hand, Fig. 5.12c shows that W-SWI-C with a Hold allocation is better than fixed period allocation with 10% multicast, see Section 5.4. This is due to the fact that the hold-release mechanism eliminates the reallocation delay. Fig. 5.12c also shows that, as the allocation for fix-period is increased, performance starts to decay because of the inflexibility of resources time scheduling.

On the other hand, Fig. 5.12b demonstrates the effect of increasing the number of masters in W-SWI-C: Hold under 10% multicast. Clearly, the performance in general improves as the number of Zen-neck surface-wave interconnects (SWIs) is increased. However, the W-SWI-C with two masters and two physical channels (SWI:2) seems to slightly outperform the SWI:3. This could be because the increase in SWI channels will increase contention on the shared medium and the arbitration delay might impact on the performance. Thus, the optimum number of masters is not always the highest. In this work, the SWI:4 is chosen as the design parameter since it offers the best performance/cost trade-off, in addition to the fact that it copes with the frequency limit for the 45nm technology.

### 5.5.2 Power Reduction

This section presents the evaluation results of an important cost metric for future interconnects, which is power consumption. The router's static and dynamic power is calculated using Orion 2.0 [122] area and power models. The modelled baseline router power has been calibrated to match the reported power measurement of the implemented SCC NoC [5] by adjusting the link activity variable in Orion 2.0. Power dissipation for wire links is calculated for the horizontal links (3.6mm) and the vertical links (5.2mm), according to the SCC measurements [5]. The transceiver power consumption projection [18, 98] is used, which is calculated to be 24mW per sub-channel. The SWI power dissipation is also calculated based on the analytical model introduced previously [34]. The GMA was designed using Verilog and then synthesized using the Synopsys Design Compiler and mapped onto the PDK 45nm technology library to calculate its dynamic power (4.8mW) and leakage power (59.3μW). Then all these values were used in our cycle accurate simulator to calculate the overall NoC power consumption.

Fig. 5.13 shows the ratio of the Mesh power consumption over the power consumption of the W-SWI-C at  $PIR = 2 \times ZLL$  for different NoC sizes, synthetic traffics and percentages of 1-to-M. Significant improvements in the NoC power consumption reduction ratio are demonstrated. For instance, the power ratio of Mesh to W-SWI-C starts from more than double ( $\sim 2X$ ) and increases up to  $\sim 10X$  as the NoC size and the broadcast percentages are increased.

However, less improvement appears in the case of multicast. This is because the multicast group members are fewer, which reduces



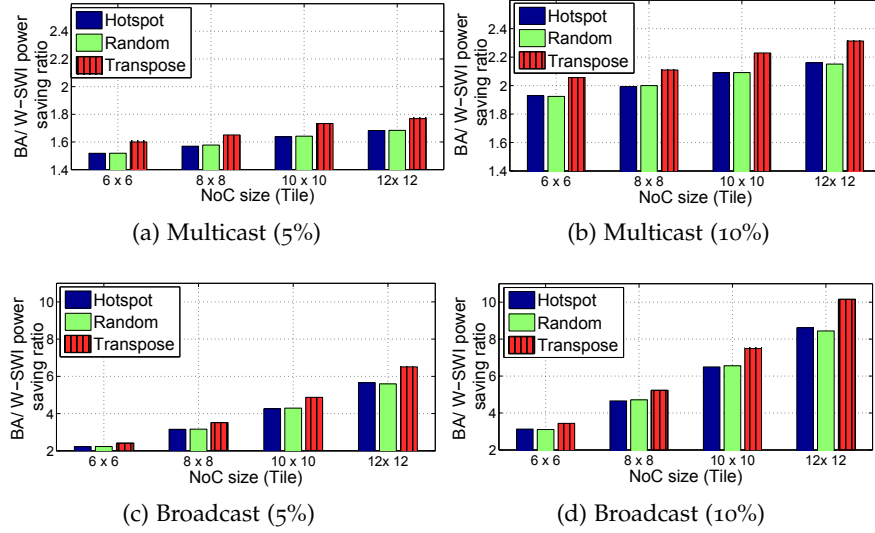
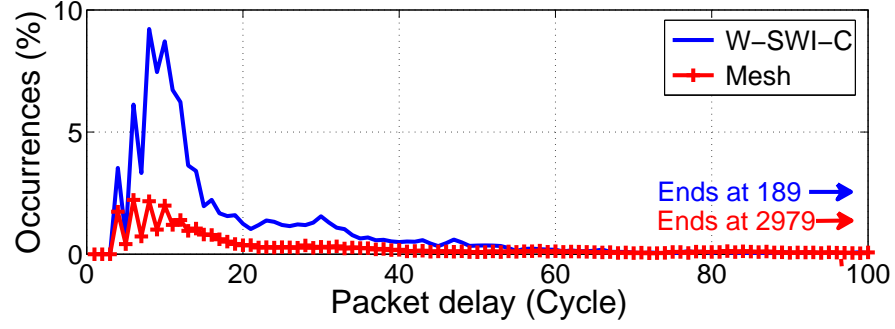


Figure 5.13: Communication power saving ratio of the W-SWI over the Mesh for different network sizes, types of synthetic traffic and broadcast/multicast ratios.

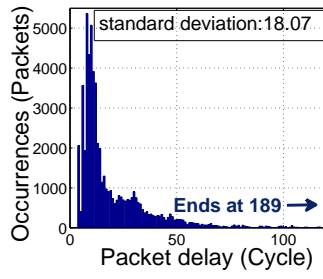
the utilization of the SWI fanout feature. Nonetheless, it still shows remarkable improvements, increasing from  $\sim 1.5X$  to  $\sim 2.3X$  proportionally to NoC size and 1-to-M ratio. In general, these findings prove that the W-SWI has a remarkable scalability and effectiveness in mitigating 1-to-M communication issues.

### 5.5.3 Quality Of Service (QoS)

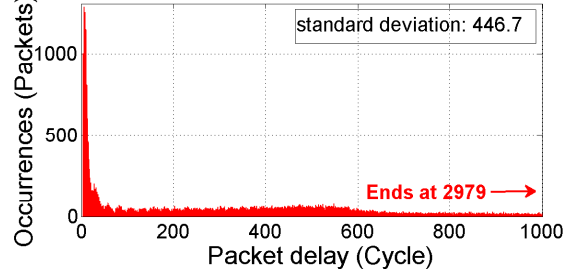
Offering guaranteed communication service is important for a range of on-chip applications that require reliable system performance. This can be measured in terms of the diversity of packet delays in passing through the interconnect fabric. Thus, this section compares the software multicast in Mesh and the W-SWI-C. Fig. 5.14 shows histograms of packet delay frequency for both architectures. There is less packet delay diversity with the W-SWI-C architecture (ranging from 4 to 189 cycles) with a standard deviation of 18.07. In contrast, Mesh shows greater packet delay diversity (ranging from 4 to 2978 cycles) with a standard deviation of 446.7. Moreover, Fig. 5.14a shows a comparison of W-SWI-C and Mesh in terms of the percentage of packet delay occurrences, which is the number of packets with a certain delay divided by the total number of packets. Unlike with the W-SWI-C, most of the curve of packet occurrence percentage of the Mesh lies under 2%. These levels of performance for Mesh will not be sufficient for applications that require highly guaranteed quality of service and high real-time performance (such as SNN) since the system behaviour and performance are unpredictable.



(a) Occurrences percentage comparison



(b) W-SWI



(c) Mesh

Figure 5.14: Packet delay distribution comparison of W-SWI and Mesh with software multicasting for a  $6 \times 4$  NoC with 5% broadcast.

#### 5.5.4 Comparison with Related Work

In this chapter, one of the state-of-the-art wire-based NoCs with a tree-based multicast scheme was replicated in order to compare it with the proposed architecture. This scheme is the VCT [15]. It has been chosen because of its efficiency and simplicity, where a minimum of modifications to the baseline router microarchitecture are required. These modifications mainly include the VCT table and the control circuits for the forking of one flit/cycle. These features will provide a fair comparison with the proposed architecture.

This scheme is based on assigning one of the virtual circuit table entries in each router to a multicast group with a unique source. Then, packet forking and routing is conducted according to the virtual circuit table. This look-up-table needs a set-up stage to define its content for each new multicast group introduced to the NoC. The set-up stage uses software multicast and the table entry is cumulatively set up. Thus, the authors acknowledged that interconnect performance is proportional to the VCT table size in the router and inversely proportional to the number of unique multicast groups injected to the NoC [15].

A big limitation of the VCT is its inability to handle wormhole traffic due to deadlock occurrence. Therefore, the chosen packet sizes were one and four flits to give a fair demonstration of the NoC's performance compression under packet-switching and virtual-cut-through-switching respectively. In addition, the VCT is limited to a turn-

Table 5.1: Average delay and PIR at the edge of NoC saturation comparison of the W-SWI and VCT.

Packet size (flit)	Multicast (%)	Average delay (Cycle)		PIR	Improvements (VCT/W-SWI)
		W-SWI	VCT		
1	5	16.74	29.65	0.315	1.77
	10	16.68	28.87	0.18	1.73
	15	16.61	29.07	0.125	1.75
4	5	16.86	96.71	0.065	5.73
	10	16.33	88.75	0.04	5.43
	15	16.6	90.5	0.03	5.45

model routing with no path diversity. Otherwise, deadlock problems could appear because diversity might introduce cycles when building the multicast tree in the set-up stage. In contrast, path diversity is a favourable feature in our architecture and has been tackled by using odd-even routing, see Section 5.3. Therefore, XY routing was used in all of the simulations in this section in order to provide a fair comparison. Moreover, according to Jerger *et al.* [15], a VCT with 512 entries per source offers good performance/cost levels in most cases. Therefore, a VCT with 512 entries was considered in all our evaluations.

Table 5.1 shows a performance comparison of the W-SWI-C and the VCT architectures with different multicast ratios. The multicast source and group members were selected randomly. The average delay and PIR is reported for the NoC saturation edge where the average delay is double the ZLL. In this PIR, the proposed architecture shows steady improvements of around  $\sim 1.8X$  and  $\sim 5.5X$  for one and four flit packet sizes respectively. Although the proposed architecture performs better with wormhole or virtual-cut-through switching, its performance is still almost double that of the VCT under basic packet switching.

#### 5.5.5 W-SWI-C for Spiking Neural Network (SNN)

The SNN modelling applications are an extreme case with 100% multicast ratio that might benefit from W-SWI. The adopted neural network model was biologically-inspired and generated from anatomical connection statistics [162]. It was assumed that each tile processes 1000 neurons along with their interconnects. These neurons were power-based mapped to the tiles based on the annealing algorithm [163]. A compression technique that aggregates neuron spikes with the same destinations, similar to the one proposed by Carrillo and *et al.* [153]) was assumed. Then, the benchmark was built after calculating the PIR

for all the mapped neurons in each tile along with their source/destination(s) for each traffic scenario.

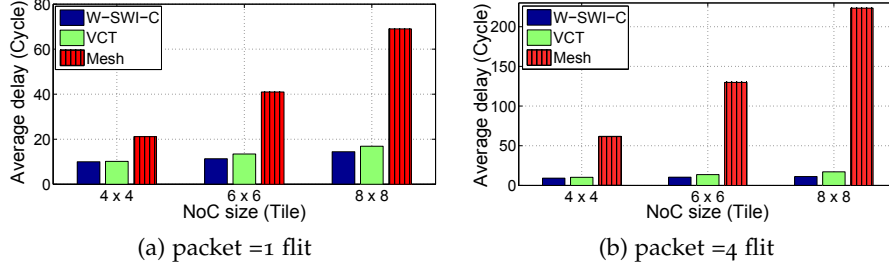


Figure 5.15: Average delay comparison of NoC ( $6 \times 4$ ) between W-SWI and VCT and Mesh under SNN benchmarks for different NoC size.

Fig. 5.15 shows an average delay comparison of the W-SWI-C, VCT and Mesh. Once more, the packet sizes were selected to be one and four to give a fair comparison with the VCT. Even though the traffic injection was relatively low with a spike size of 16 bits and the number of unique multicast groups is relatively small, the proposed architecture shows better performance than the VCT and huge improvements over Mesh, especially for a packet size of 4 flits. Moreover, the improvement increases from  $\sim 2\%$  to  $\sim 30\%$  over VCT and from  $\sim 46\%$  to  $\sim 95\%$  over Mesh when the NoC size is increased. This is due to the efficient handling of multicast traffic with high graph density and graph degree. Thus, the W-SWI is more scalable than the VCT and Mesh for SNN modelling architectures.

#### 5.5.6 Area Overhead Evaluation

It is essential to evaluate chip area overheads for the extra on-chip circuits required for the proposed architecture. Firstly, it is assumed that the active area calculated for transceivers in previous research [18] is the only part scaled down when moving to 45nm technology, while the passive parts remain almost the same since they are proportional to the channels' operational frequency range. Therefore, the projected transmitter area is  $4870\mu\text{m}^2$  per sub-channel, while the projected receiver area is  $260\mu\text{m}^2$  per sub-channel, where the active area is proportional to the square of the scaling factor [148].

Secondly, the area of the extra router port (buffer, crossbar and related circuits) is calculated using the Orion 2.0 [122] model as  $0.427\text{mm}^2$ . The modelled baseline router area is a 6% less than the reported implemented router area [5], which is acceptable for the purpose of comparison evaluation in this study. In addition, the GMA was designed using Verilog and then synthesized using the Synopsys design compiler and mapped onto the PDK 45nm technology library to calculate its area ( $0.0114\text{mm}^2$ ), to which the Tx/Rx estimated area of  $0.0438\text{mm}^2$  was added. Likewise, the VCT with a 512 entry/source

Table 5.2: Area overhead evaluation for W-SWI-C, proposed W-SWI-D and VCT-512 [15] over baseline architecture (Mesh).

NoC component		Area per item ( $\text{mm}^2$ ) for 45nm technology			
Component	No.	Mesh	W-SWI-C	VCT	RF-I
Router	24	1.0853	1.5124	1.0853	1.5124
Transmitter	4	-	0.1558	-	0.1558
Receiver	24	-	0.0083	-	0.0083
Global arbiter	1	-	0.0552	-	0.0552
VCT table	24	-	-	2.3608	-
Wire Links	1	13.653	13.653	13.653	13.653
Total extra area over Mesh= (all components $\times$ their No.)- Mesh area ( $\text{mm}^2$ )			11.13	56.66	11.9
NoC/SCC-die area (%)	7		8.96	16.99	9.1

lookup table was designed using Verilog and then synthesised. On the other hand, to compare other emerging interconnects, the RF-I's transmission line area was calculated and considered to be routed through the chip (NoC size  $6 \times 4$ ) as a U shape passing through all nodes [18]. In addition, a transmission line with a pitch of  $12\mu\text{m}$  was considered in calculations of its area overhead.

Table 5.2 shows area considerations for the Mesh, proposed W-SWI-C, VCT and RF-I. Obviously, most of the extra area overhead for the W-SWI-C, of around 1.96% is due to the extra router port. Moreover, the W-SWI-C offers a better die area-performance trade-off compared to RF-I transmission lines that offer the same connectivity [18], since fat transmission lines need to be implemented through the chip. Not only that, but the W-SWI-C also beats the VCT in area overhead at around five times less. Therefore, the W-SWI-C succeeds these architectures in terms of low area overheads and circuit complexity.

## 5.6 SUMMARY AND CONCLUSION

In this chapter a hybrid wire-SWI architecture has been proposed for on-chip communication that is able to tackle 1-to-M traffic efficiently and gives an attractive area/performance trade-off. Zenneck surface wave low power dissipation, high signal propagation speed and fanout capability all help to significantly mitigate the 1-to-M communication issues that the NoC-based CMP in particular suffers from. In addition, centralized (W-SWI-C) arbitration and allocation techniques along with a multicast routing scheme for this architecture are proposed and dis-

cussed. The evaluation results show significant improvements in terms of average delay, saturated PIR and power consumption with a relatively small die area penalty compared to state-of-the-art-architectures. In general, the results demonstrate the high scalability of the W-SWI-C for future NoC-based CMPs. Next chapter will try to solve the SWI multicast contentions issues in decentralized manner.

## WIRE AND SURFACE-WAVE ARCHITECTURE WITH DECENTRALIZED CONTROL FOR MULTICAST

---

### 6.1 INTRODUCTION AND MOTIVATION

Continuous increase of the number of cores in CMPs enabled by the CMOS technology scaling will bring us soon to the many-core era. 1000's of cores are foreseen in the near future. Therefore, the number and size of communication between cores are expected to increase exponentially. However, CMP performance and power consumption depend both on NoCs and cache coherence protocols.

Cache coherence protocols depend on a range of 1-to-M communication patterns such as multicast invalidation requests (directory-based protocols) and broadcasting ordering tokens (broadcast-based protocols) [15, 16]. This could be catastrophic in terms of performance and power budget since it leads to high congestion, hotspots, and power consumption, as mentioned in the last chapter.

On the other hand, regular wire projected issues [7] inspired many researchers to propose hybrid NoC architectures. These promising paradigm could solve future many-core scaling issues such as traffic load increase, global communication and multicast communication. These hybrid NoC architectures utilize a set of alternative merging interconnects, such as RF interconnects RF-I [18], WiNoC [22] and optical interconnects [28]. However, the main issue facing most of these promising hybrid architectures is the contentions [23]. In the previous chapter these issues were tackled using a dead-lock free centralized approach that is more superior for broadcast than the general case of multicast. In this chapter we further explore the design of hybrid W-SWI to efficiently resolve the issues of multicast contentions in a decentralized manner. The major contributions of this chapter are:

- A novel technique called stretched multicast is proposed maximizing the utilization of S-SWI for on-chip multicast.
- A novel decentralized arbitration scheme is proposed which can maximize the slack-time scheduling in a multiple-source and multiple-destination scenario.
- It has been shown that the proposed arbitration and routing method is deadlock-free using a newly proposed ID-tagging scheme.
- The proposed architecture is rigorously evaluated. Comparing to the previous approaches, it can achieve an improvement for

Table 6.1: Notations used in this chapter.

$M_x$	Master node number $x$
$S_x$	Slave node number $x$
$P_t(S_x)$	Probability that $S_x$ is free at time $t$
$N$	NoC size
$Rq_{x,y}$	Request from $M_x$ to $S_y$
$N_m$	Number of masters
$MG_i$	Set of multicast group members
$N_v$	Number of VCs
$Tag$	ID-tagging for flow control
$F$	Flit
$ST$	Time when the multicast flit is delivered to all its destinations
$RT$	Reservation table
$RV_x$	Request vector from $M_x$
$T_x$	Time slot sequence $x$
$Gr_{x,y}$	Grant signal from $S_y$ to $M_x$
$v_i$	Input VC
$v_o$	Output VC

power reduction and communication speed up to 63% and 12X, respectively.

## 6.2 DECENTRALIZED ARBITRATION AND ALLOCATION

In the centralized approach, a master has to wait until all the requested slaves (multicast group) are free in order to avoid deadlock, as mentioned in Chapter 5. To mathematically express the problem in this approach, assume a master ( $M_i$ ) is sending a request vector ( $RV_i$ ) to the GMA to allocate a set of multicast-group ( $MG_i$ )  $\subset \{S_0, \dots, S_{N-1}\}$ , where  $S_j$  is a slave, and  $N$  is the NoC size. The probability that a slave ( $S_j \in MG_i$ ) is free in time slot  $t$  is denoted by  $P_t(S_j)$ . Therefore, the probability that the  $M_i$  request ( $Rq_{i,j}$ ) will be granted is the intersection probability ( $P_t(S_x) * P_t(S_y) \dots * P_t(S_z)$ ), which is clearly less or equal than  $P_t(S_j)$  for all  $S_j \in MG_i$ . This will keep free requested slaves idle until all the members of  $MG_i$  are free. Moreover, GMA might be considered as a single point of failure, which makes the architecture less robust to any possible errors. Therefore, a less constrained and centralized arbitration and allocation technique are required to seize any available opportunity to utilize the physical channels of SWI. Table 6.1 shows a list of notations used in this chapter.



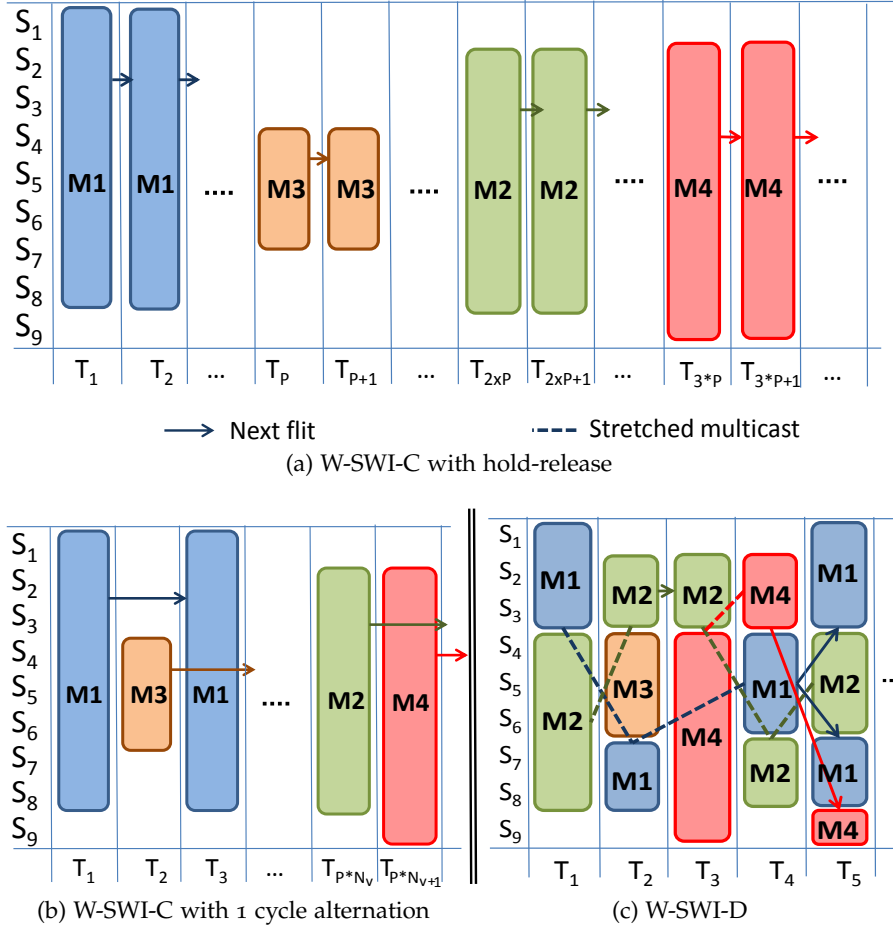


Figure 6.1: Example shows traffic flows from four masters multiplexed in SWI where (a) W-SWI-C architecture with Hold allocation mechanism (b) W-SWI-C architecture with fixed period (one cycle) alternation and two VSWI, and (c) W-SWI-D architecture (decentralized arbitration and stretched multicast), where M is a master, S is a slave and T is a time slot.

### 6.2.1 Stretched Multicast

To improve SWI utilization and avoid centralized arbitration, this section proposes an alternative technique we refer to as a stretched multicast. This approach allows a master to transmit a multicast flow to any available slaves. Although this technique requires partial retransmission, it allows the concurrent execution of several overlapping multicast communications. This is achieved by enabling a master to transmit multicast flow to any number of currently free slaves and to retransmit it to the rest later. Consequently, the decision should be determined at the slave end in a decentralized arbitration. This can be realized by any simple independent fair arbiter in each slave. Round-robin has been chosen since it provides stronger fairness than rotating or random arbiters and requires less circuit complexity than matrix and queuing arbiters [51, 57].

There are many possible scenarios where this scheme will show better contention handling, fairness, and higher SWI utilization. For instance, Fig. 6.1 shows a comparison of contention handling of W-SWI-C and a decentralized architecture, we refer to it shortly as W-SWI decentralized (W-SWI-D). Fig. 6.1a shows W-SWI-C with hold-release mechanism (Hold) where master hold the physical SWI channel until his request is fulfilled, see Chapter 5. On the other hand, Fig. 6.1b shows W-SWI-C with 1 cycle allocation period for each master that already reserved one of two VSWIs. Clearly, even though we assume that the W-SWI-C in both cases manages to allocate all members of  $MG_i$  at  $T_1$ , it offers less fairness and utilization of the SWI links. In addition the multiplexing between flows is limited to the packet level. In contrast, although the same flit might be retransmitted (up to  $N_M$ , where  $N_M - 1$  is the number of master nodes), the stretched multicast offers higher fairness and utilization of the SWI in this case, as shown in Fig. 6.1c. This is because the reallocation of slaves on flit bases prevents flows from blocking each other. In contrast, the blocking period ( $Time_{block}$ ) in W-SWI-C with Hold mechanism can be predicted to be up to:

$$Time_{block} = (Packet_{size} \times N_m) - 1 + Time_{congestion} \quad (6.1)$$

while the blocking period with W-SWI-C with a fixed period alternation could be up to:

$$Time_{block} = (Packet_{size} \times \frac{N_m - N_v}{N_m - N_v}) + Time_{congestion} \quad (6.2)$$

where  $N_v$  is the number of VC,  $Time_{congestion}$  is the duration of any congestion in the master (the channel owner) or the slave that would cause idle time slots. However,  $Time_{congestion}$  might equal zero if there is no heavy load at either ends. Thus, the stretched multicast improves fairness by reducing the blocking period between flows. However, partial allocation of multicast group members in W-SWI-D might cause a multicast dependency that leads to a deadlock scenario.

### 6.2.2 Deadlock-free Flow Control

In section 5.4.2, the use of wormhole switching in blocking channels limits the transmission by constraining it. The constraint was allocating all  $MG_i$  to avoid multicast dependency. In order to break multicast dependencies without the need to allocate all the requested slaves ( $MG_i$ ) at a given time slot as suggested by Section 6.2.1, each master should have its own virtual non-blocking channels for every slave. This will allow wormhole switching at packet level but with cut-

through switching at flit level. That means each master should have its own statically allocated virtual link. Therefore, since each master already has its own physical channel frequency, virtually non-blocking channels can be achieved at the slave router by using a statically allocated VC where each master transmits via one VC ( $N_v = N_M$ ).

The other design option is the proposed ID-tagging-based flow control (Tag). This is simply achieved by tagging each flit (F) with the transmitter master's ID ( $M_x$ ) so that  $\text{Tag} = M_x$ , where  $\text{Tag}, M_x \in \{0, \dots, N_M - 1\}$ . Then, at the reservation table (RT), the allocation entry is distinguished based on the input buffer (i), Tag and the input VC ( $V_i$ , if the design include VSWI). For simplicity, the ID-tagging remains unchanged in the router output port while being drained for the local PE.

Fig. 6.2 shows an example of a router micro-architecture providing two virtual non-blocking channels by using either ID-tagging or statically allocated VCs. Obviously, ID-tagging allows designers to choose a virtual-channelless design or at least not constrain the number of VC to be  $N_v \geq N_M$ . Consequently, the router would require less area and power overheads. For this example, if alterations are limited to the router port linked to SWI, then our calculations estimate a reduction in area of  $\sim 2\%$ , and in power of  $\sim 5\%$ . Moreover, ID-tagging gives the freedom to use the VC for other purposes such as multiplexing flows from the same master (VSWI) in cases of congestion.

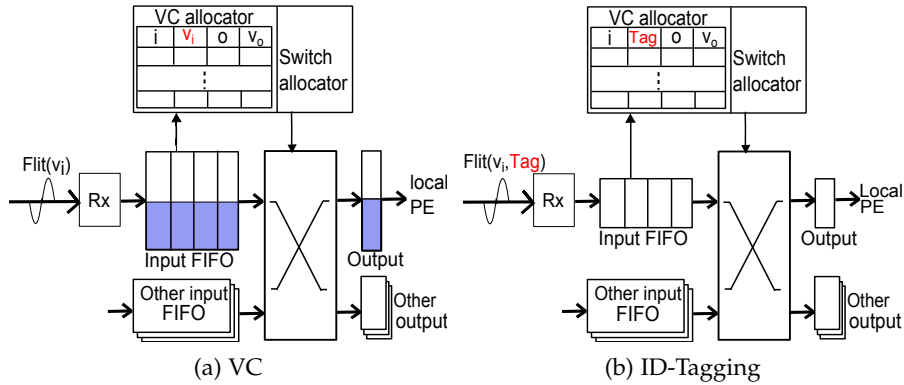


Figure 6.2: Illustration of router micro-architecture with ID-Tagging based flow control and VC flow control.

To prove that this scheme breaks the multicast dependencies on the SWI layer, assume that we have requests ( $Rq_{i,j} \in RV_i$  for a  $M_i$  to any slave ( $S_j \in \{1, \dots, N\}$ ). These requests should be granted within a finite time ( $T_f$ ). This is to prove, that for all  $Rq_{i,j} = 1$ , then  $Gr_{i,j} = 1$  within  $T_f$ , where  $Gr_{i,j}$  is a grant signal. The probability that a slave local arbiter (RR) grants a master's request at the next time slot  $T_x$  is  $P_{T_x}(S_j) = \frac{1}{N_M}$ . In the worst case scenario,  $M_i$  has just been granted in  $T_{x-1}$  and all masters are requesting  $S_j$  at  $T_x$ . Thus,  $M_i$  has to wait until the  $S_j$  RR arbiter grants all other masters. Assuming the average

Table 6.2: Comparison of highlighted features of the centralized and decentralized approaches for the proposed architecture.

Features	Centralized approach (W-SWI-C)	Decentralized approach (W-SWI-D)
Flit retransmission	none	up to $N_M - 1$
Probability of transmission	$P_t(S_x) * P_t(S_y) \dots * P_t(S_z)$	$P_t(S_j)$
Flow control	wormhole	wormhole on packet level and packet-switching on the flit level
VSWI utilization	on the slave end	on both master and slave end
Channel establishment	per packet	per flit

delay to serve each master request is  $T_D$ . Therefore, the maximum waiting period is  $= N_M \times T_D$  and thus  $P_{T_x + N_M \times T_D}(j) = 1$ . Therefore, the serving time  $ST_i$  needed for a flit to be delivered to all slaves and ejected from the output buffer in  $M_i$  is:

$$ST_i = \max \left\{ \bigcup_{T_x = T_0 \rightarrow \max(T_x)} \{T_x \times (Rq_{i,0}, \dots, Rq_{i,N-1})\} \right\} \quad (6.3)$$

where the maximum period for delivering each flit per slave is  $= T_x + N_M \times T_D$  for all  $Rq_{i,j}$ , and  $j \in \{0, \dots, N-1\}$ , and  $T_0$  is the request insertion time. As a result,  $1 \leq ST_i \leq N_M \times T_D$ . For instance, assume that  $M_1$  send requests  $\{Rq_{1,1}, Rq_{1,2}, Rq_{1,3}\}$  for slaves  $\{S_1, S_2, S_3\}$  at  $T_0$ . If each request has been granted in time slot  $\{T_3, T_3, T_1\}$  respectively then, according to Equation. 6.3,  $ST_1 = T_3$ . In other words, the maximum serving time for the whole multicast is the maximum serving time among all the multicast group members, which is a finite time.

Table 6.2 highlights a comparison of the main features of the centralized and decentralized approaches. These features will rationalize some of evaluation results in section 6.3.

### 6.2.3 Communication Protocol

The communication protocol for the decentralized approach has fewer control signal types than that in the centralized approach, but with slight composite interface procedure. Algorithm 6.1 and Algorithm 6.2 show the procedure used for master and slave interfaces, respectively. First at master end, a master ( $M_i$ ) interface sends a  $RV_i$  via SWI that consists of a request ( $Rq_{i,j}$ ) to each slave ( $S_j$ ) wherever  $MAB(j) = 1$ ;

---

**Algorithm 6.1** Master interfaces procedure to communicate via SWI.

---

**Define:** -

$i$  :local node ID

**Input:** -

$P_{in}$  :input port

MAB :multicast-address-bits

**Output:** -

$RV_i$  :Request vector

```

1: for all  $P_{in} \in \{N, E, S, W, L\}$  do
2:   if Reserve( $P_{in}, MAB$ ) == SW then
3:     for all  $j \in \{0, \dots, N-1\}$  do
4:        $RV_i[j] = RV_i[j] \vee MAB[j]$ ;
5:     end for
6:   end if
7: end for
8: for all  $j \in \{0, \dots, N-1\}$  IF any Rx( $Gr[j]$ ) == 1 do
9:   Tx(Flit = FIFO());
10:   $RV_i[j] = MAB[j] = 0$ ;
11: end for
12: for all  $j \in \{0, \dots, N-1\}$  IF all MAB[j] == 0 do
13:   FIFO.eject;
14: end for
15: Tx( $RV_i$ );

```

---

see lines 1 to 7 in Algorithm 6.1. The master interface keeps track of which slave has been served by updating an instant of MAB of the current flit; see lines 8 to 11 in Algorithm 6.1. Then, if all  $MG_i$  members have received the multicast flit, it will be ejected and a new set of requests might be sent based on the next flit MAB; see lines 12 to 14 Algorithm 6.1.

---

**Algorithm 6.2** slave interfaces procedure to communicate via SWI.

---

**Define:** -

$j$  : local node ID

**Input:** -

$RV_0[j], \dots, RV_{N_M-1}[j]$  :masters requests

**Output:** -

$G_{0,j}, \dots, G_{N_M-1,j}$  : masters grants

```

1: for  $i = 0 \rightarrow N_M - 1$  do
2:    $R_i = Rx(RV_i[j]);$ 
3: end for
4: for all  $i \in \{0, \dots, N_M - 1\}$  IF all  $Gr_{i,j} == 0$  do
5:    $i = RR(R_0, \dots, R_{N_M-1})$  ;
6:    $Tx(Gr_{i,j} = 1);$ 
7: end for
8: if  $Rx(Flit) == 1$  then
9:    $Tx(G_{i,j} = 0)$  ;
10: end if
```

---

At the slave end, the local RR arbiter will determine which master it will listen to in the next time slot and send the  $Gr_{i,j}$  signal to, see lines 4 to 7 in Algorithm 6.2. Then when the flit has been received the grant signal is disabled and the arbitration process restarted. Both  $Rq_{i,j}$  and  $Gr_{i,j}$  are transmitted via SWI using On-off keying (OOK) modelling for simplicity. When the RR grants the request, the data handshake phase starts by sending the data flit to all slaves who responded to the request and waits for them to acknowledge reception before resetting the requests intended for them. The adopted handshaking protocol is a non-return-to-zero protocol [164]. Then, if all  $MG_i$  members have received the multicast flit, it will be ejected and a new set of requests might be sent based on the next flit MAB.

### 6.3 SYSTEM LEVEL EVALUATION AND DISCUSSION

This section presents results obtained from our cycle-accurate NoC simulator which was built by modifying the existing Noxim simulator [121] for the W-SWI-C, W-SWI-D, VCT and the Mesh. In this thesis, the

Intel SCC [5] tile specification is adopted as the baseline architecture, as mentioned earlier. Fig. 6.3 presents the procedure and simulation tools used in this section to generate the benchmarks, simulate the proposed architecture along with with related state-of-the-art architectures, and synthesize the extra proposed components. These all then used in the NoCs cycle accurate simulator to obtain the system-level evaluation results.

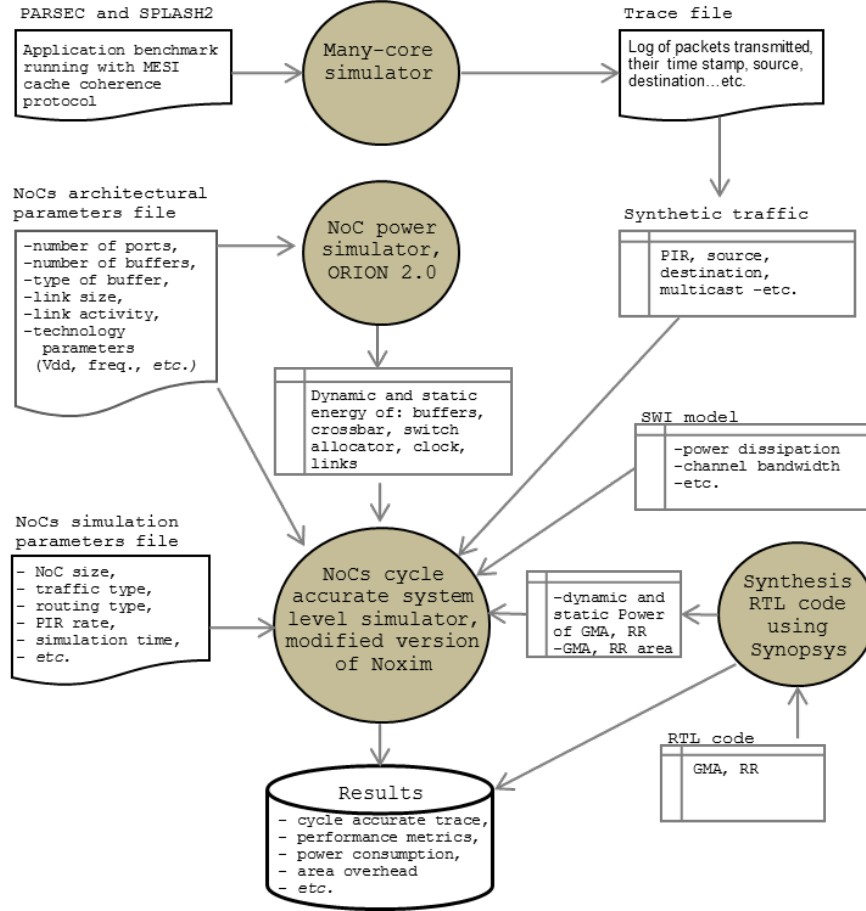


Figure 6.3: Simulation flow to obtain the results.

Packet sizes of 1, 4 and 12 flits were chosen as an example to demonstrate the behaviour of the proposed architecture under packet-switching, virtual-cut-through-switching and wormhole-switching flow control, respectively. The number of master nodes is based on the available frequency range for 45nm, which was estimated to be four channels (plus the frequencies specified for control signals). However, this frequency range is scaling with technology [18]. As a result, the number of master nodes was increased when simulating larger NoCs, assuming that the technology will have been scaled too. In addition, a VSWI number of four was chosen, which realizes a better performance/cost trade-off. In addition, in the evaluation in this section

the VC was chosen to be equal to the VSWI for simplicity of router architecture.

The simulation was conducted with synthetic traffics, which are: (1) Random, where packets are transmitted randomly with uniform probability to other nodes; (2) Hotspot, which is the same as the random but with specific nodes called hotspots, four in this case, with higher probability of traffic dispatched to them; (3) Transpose, where a node sends a packet to other node that has its address transposed. These synthetic traffics adjusted to inject a specific percentage of broadcast (1-to-all) or multicast (1-to-M). The source nodes of this multicast traffic are selected randomly during the simulation, while the rest of the traffic consist of normal unicast packets according to the named synthetic traffic. In addition, in the case of multicast, the destinations of these packets are also selected randomly. In addition, the evaluation of the proposed architecture and baseline architecture includes real application benchmarks whose details are shown in Section 6.3.3.

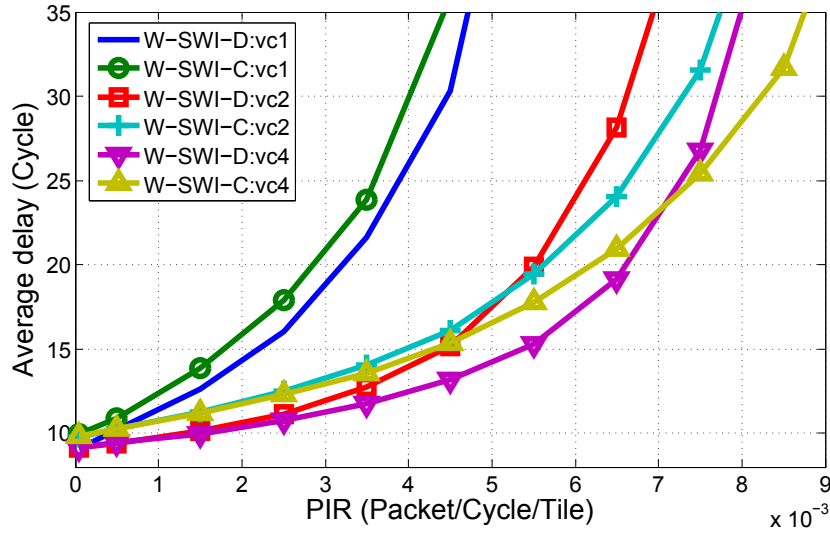
#### 6.3.1 Performance Improvements

This section presents performance evaluation for both of the proposed architectures W-SWI-D and W-SWI-C with multicast communication. For instance, Fig. 6.4 shows comparison between W-SWI-C and the proposed decentralized arbitration and routing scheme (W-SWI-D) for different VC sizes and NoC sizes. Obviously, the proposed W-SWI-D shows better performance than W-SWI-C with one VC. Improvements are ranging  $\sim 10\%$  to  $\sim 12\%$  at double the ZLL. This is due to the fact that with one VC, ID-tagging allows each master to have a virtually non-blocking channel.

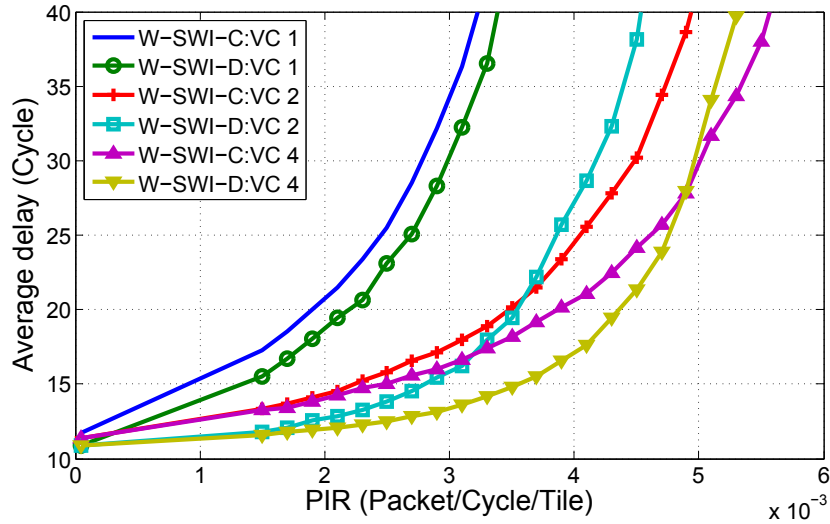
Moreover, performance results show that even for higher  $N_v$ , the W-SWI-D is better before the NoC saturation. These improvements are increasing as we increase the NoC size. For instance, as Fig. 6.4 shows at  $PIR = 1.5 \times ZLL$  the improvements with two VC are  $\sim 6\%$ ,  $\sim 13\%$ ,  $\sim 15\%$  for NoC size  $6 \times 6$ ,  $8 \times 8$  and  $10 \times 10$ , respectively. In addition, the gap is increased as the VC number is increased. Thus, the improvements with four VCs are  $\sim 15\%$ ,  $\sim 17\%$ ,  $\sim 21\%$  for NoC size  $6 \times 6$ ,  $8 \times 8$  and  $10 \times 10$ , respectively. These improvements are the result of higher SWI utilization, which is up to 2.5% increase in the average flit number transmitted via SWI, as has been observed from the simulation. However, at the saturation point the multiplexing delay between masters starts to overcome the improvements in the SWI utilization. Thus, this architecture is favourable for low or medium load applications.

However, the decentralized approach is not suitable for scenarios with broadcast (1-to-all) traffic. This is due to the fact that broadcast traffic does not benefit from the flexibility offered by the stretch-multicast and the ability to have two or more simultaneous broadcasts

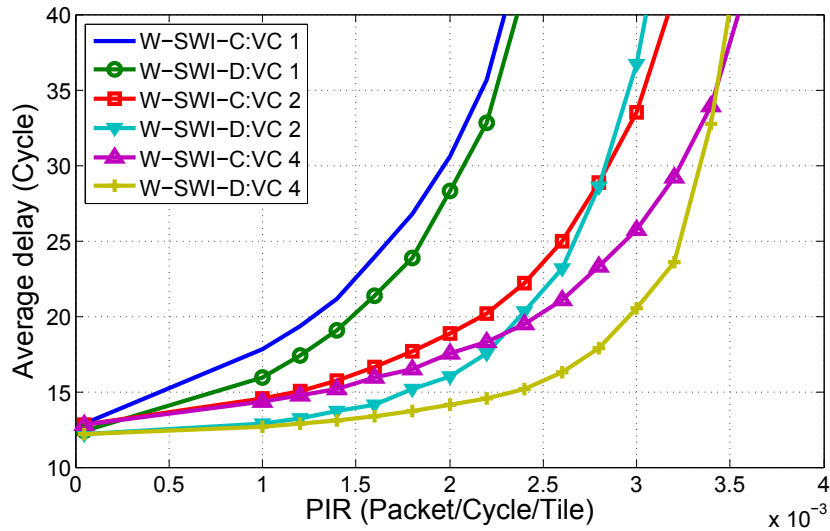




(a) 6x6 NoC

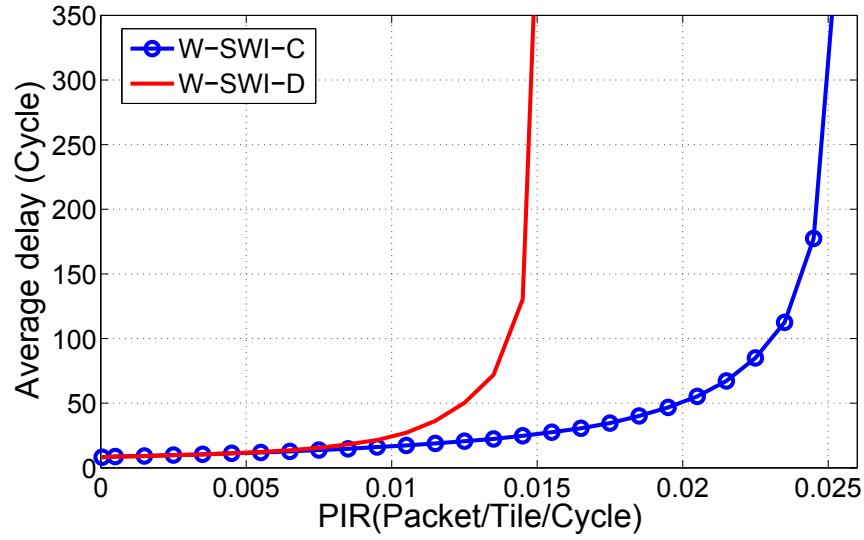


(b) 8x8 NoC

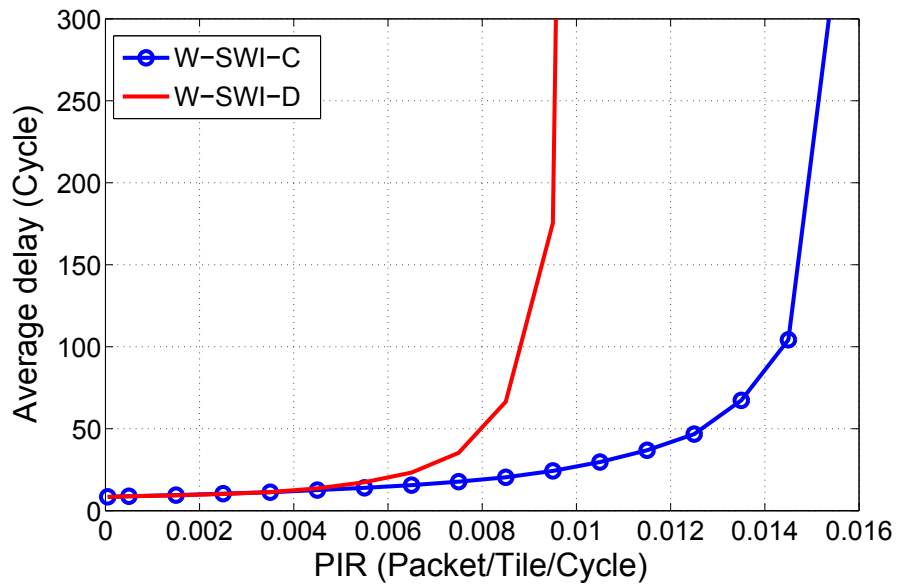


(c) 10x10 NoC

Figure 6.4: Comparison between the average delay of W-SWI-C and W-SWI-D under uniform synthetic traffic with 10% multicast ratio for different VC number and NoC size.



(a) 5% Broadcast



(b) 10% Broadcast

Figure 6.5: Comparison between the average delay of W-SWI-C and W-SWI-D under uniform synthetic traffic with 5% and 10% broadcast ratios.

unlike the case of multicast. Therefore, it ends up with adding the delay of channel establishments per flit for the overall latency and degrade the system performance. This is clearly shown in Fig. 6.5 where the W-SWI-D is saturated much faster than the W-SWI-C when the traffic consists of broadcast traffic of such percentages: %5 and %10.

In addition, regarding improvements over the baseline architecture, W-SWI-D improvements ratio over the Mesh compared to W-SWI-C seems to be almost steady (up to 2X), even when the NoC size increased as shown by the Table 6.3. These results are obtained with PIR that is  $1.5 \times \text{ZLL}$  and with 10% multicast traffic ratio.

### 6.3.2 Power Reduction

This section presents the evaluation results of the power consumption of the proposed decentralized architecture compared to centralized and baseline architectures. The router's static and dynamic power is calculated using the Orion 2.0 [122] model. The modelled baseline router power has been calibrated to match the reported power measurement of the implemented NoC [5]. In addition, power dissipation for wire links is calculated for the X (3.6mm) and the Y (5.2mm) directions. The transceiver (Tx/Rx) power consumption projection [18] is used (24mW per sub-channel). SWI power dissipation is also modelled based on the analytical model introduced previously [34]. The GMA [36] and the local RR arbiter was designed using Verilog and then synthesized using the Synopsys Design Compiler and mapped onto the PDK 45nm technology library to calculate its dynamic power (4.8mW) and leakage power (59.3μW). Then all these values were used in our cycle simulator that based on Noxim to calculate the overall NoC power consumption. Table 6.3 shows the power consumption saving ratios for both W-SWI-C and the proposed W-SWI-D over the Mesh for different NoC sizes and  $N_v$ . Evidently, both architectures demonstrate significant improvements in the NoC power consumption reduction ratio over Mesh. In addition, even though the W-SWI-D may retransmit the same flit many times (up to  $N_M - 1$ ) via SWI it achieves power saving (1 – 2%) due to reduction of the flit delivery time at PIR where latency reaches  $1.5 \times \text{ZLL}$ . These findings prove that the W-SWI-D is more effective in mitigating 1-to-M communication issues in terms of power consumption for low or normal load.

### 6.3.3 Evaluation with Real Application Benchmark

In order to demonstrate the effectiveness of both the proposed architectures W-SWI-D and W-SWI-C for real applications, a set of application benchmarks from standard suites have been considered. These application benchmarks are PARSEC [150] and SPLASH2 [151]. These benchmarks are built based on the traffic analysis of communication

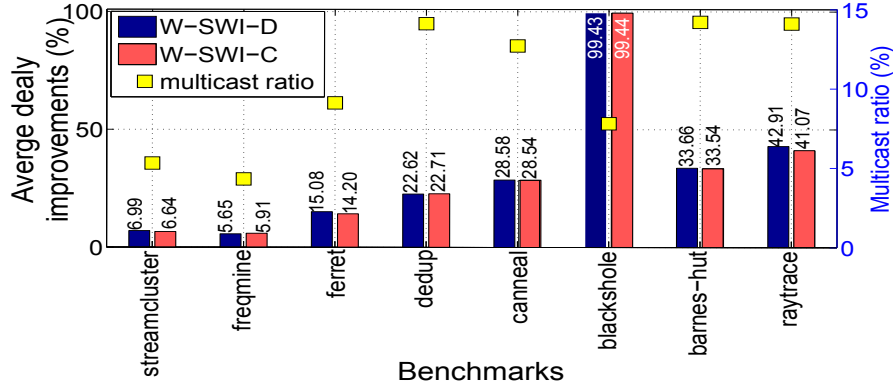
Table 6.3: Results improvements over Baseline architecture (Mesh) comparison between W-SWI-C and the proposed W-SWI-D with 10% multicast ratio.

NoC size	VC no.	PIR ( $\times 10^{-4}$ )	improvements over Mesh			
			Average delay (x)		Power (%)	
			W-SWI-C	W-SWI-D	W-SWI-C	W-SWI-D
6x6	1	15	9.64	10.61	63.34	63.64
	2	35	10.15	11.15	55.2	56.03
	4	45	10.09	11.77	53.57	55.24
8x8	1	17	8.65	9.6	56.84	57.24
	2	29	8.46	9.38	52.25	53.56
	4	33	8.61	10.58	51.41	52.94
10x10	1	12	8.24	9.16	55.48	55.77
	2	22	7.14	8.21	50.14	51.11
	4	24	7.55	9.69	49.56	50.99

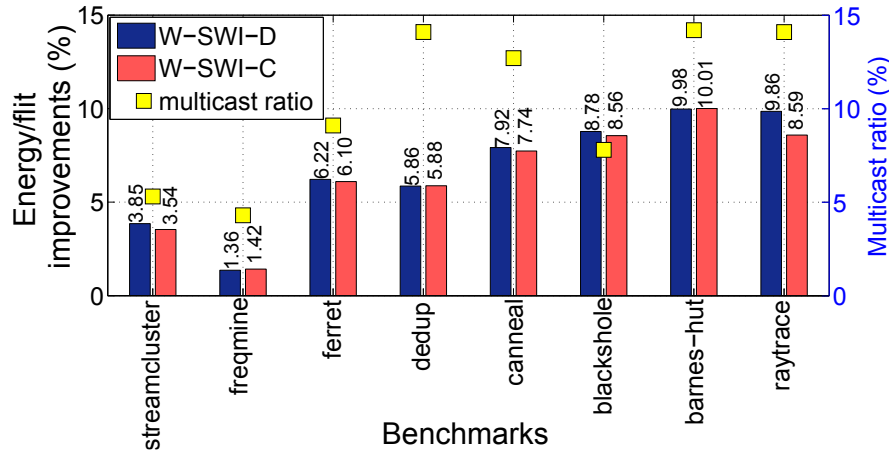
trace-files generated from the CMP simulator [75] where all the benchmark applications were run with MESI cache coherence protocol. This protocol is well-known and used in many multi-processor systems [165]. As a result, based on this traffic analysis set of synthetic traffics have been built, which have the same injection rates, packet size and source/destination(s) of each multicast and unicast traffic flows in these application benchmarks. These synthetic traffics then run for a million cycles with our cycle-accurate system-level NoC simulator.

Fig. 6.6a presents the performance improvement gained using the proposed W-SWI-C and W-SWI-D architectures over the Mesh for a NoC size of  $10 \times 8$ . In general, the average delay improvements of the W-SWI-C and W-SWI-D over Mesh are almost similar and range from  $\sim 5$  to  $\sim 99\%$ . Moreover, these improvements are clearly proportional to the percentage of the multicast's ratio from the total PIR (4 to 14.2%). An exception to this pattern is the case of the blackhole benchmark, where the improvement over Mesh is around 99%, even though the multicast ratio is 7.8%. This is due to the nature of the multicast traffic with high source hotspots that cause the Mesh to quickly become saturated. In contrast, the proposed architecture's skip-links have more ability to effectively alleviate such traffic issues.

On the other hand, Fig. 6.6b also demonstrates the average energy/flit improvements over Mesh, where a flit is 128 b as mentioned earlier. Once again, W-SWI-C and W-SWI-D achieved better rates of energy/flit over Mesh of up to  $\sim 10\%$  and the improvements are proportional to the multicast percentage. However, most of these benchmarks run with relatively low PIR. Therefore, as shown in Section 6.3.1, the W-SWI-D outperforms the W-SWI-C under low load. As a result, even though the W-SWI-D uses retransmissions that increase power con-



(a) Average delay improvements over Mesh



(b) Energy/flit improvements over Mesh

Figure 6.6: Comparison between the average delay and energy improvements of W-SWI-C and W-SWI-D over Mesh under real applications benchmarks from PARSEC [150] and SPLASH2 [151] for  $10 \times 8$  NoC.

sumption, it is better than W-SWI-C in some benchmarks in terms of power. In general, these results prove the potential of the proposed architectures for general-purpose future NoC-based CMPs.

#### 6.3.4 Area Overhead Evaluation

It is essential to evaluate chip area overheads for the extra on-chip circuits required for the proposed architecture. Firstly, it is assumed that the active area calculated for transceivers in previous research [18] is the only part scaled down when moving to 45nm technology, while the passive parts remain almost the same since they are proportional to the channels' operational frequency range. Therefore, the projected transmitter area is  $4870\mu\text{m}^2$  per sub-channel, while the projected receiver area is  $260\mu\text{m}^2$  per sub-channel, where the active area is proportional to the square of the scaling factor [148].

Table 6.4: Area overhead evaluation for W-SWI-C, proposed W-SWI-D and VCT-512 [15] over baseline architecture (Mesh).

NoC component		Area per item ( $\text{mm}^2$ ) for 45nm technology				
Component	No.	Mesh	W-SWI-C	W-SWI-D	VCT	RF-I
Router	24	1.0853	1.5124	1.5931	1.0853	1.5124
Transmitter	4	-	0.1558	0.1558	-	0.1558
Receiver	24	-	0.0083	0.0083	-	0.0083
Global arbiter	1	-	0.0552	-	-	0.0552
Local arbiter	24	-	-	0.0309	-	-
VCT table	24	-	-	-	2.3608	-
Wire Links	1	13.653	13.653	13.753	13.653	13.653
Total extra area over Mesh= (all components $\times$ their No.)- Mesh area ( $\text{mm}^2$ )			11.13	13.01	56.66	11.9
NoC/SCC-die area (%)	7		8.96	9.42	16.99	9.1

Secondly, the area of the extra router port (buffer, crossbar and related circuits) is calculated using the Orion 2.0 [122] model as  $0.427\text{mm}^2$ . The modelled baseline router area is a 6% less than the reported implemented router area [5], which is acceptable for the purpose of comparison evaluation in this study. Thirdly, the GMA (for W-SWI-C) and RR (for W-SWI-D) was designed using Verilog and then synthesized using the Synopsys design compiler and mapped onto the PDK 45nm technology library to calculate its area. Their area was found to be  $0.0114\text{mm}^2$  and  $0.0002\text{mm}^2$ , respectively, and to which the Tx/Rx (for control signals) estimated area of  $0.0438\text{mm}^2$  and  $0.0307\text{mm}^2$ , respectively, was added. Likewise, the VCT with a 512 entry/source lookup table was designed using Verilog and then synthesised. On the other hand, to compare other emerging interconnects, the RF-I's transmission line area was calculated and considered to be routed through the chip (NoC size  $6 \times 4$ ) as a U shape passing through all nodes [18]. In addition, a transmission line with a pitch of  $12\mu\text{m}$  was considered in calculations of its area overhead.

Table 6.4 shows area considerations for the Mesh, proposed W-SWI-C, W-SWI-D, VCT and RF-I. Obviously, most of the extra area overhead for the W-SWI-C and W-SWI-D, of around 1.9% is due to the extra router port. However, the W-SWI-D area overhead is higher than the W-SWI-C ( $\sim 2.3\%$ ,  $\sim 2\%$ , respectively). This is mostly due to increasing the VC allocation unit area in all routers to implement the ID-tagging scheme. Moreover, the W-SWI-C offers a better die area-performance trade-off compared to RF-I transmission lines that offer the same connectivity [18], since fat transmission lines need to be implemented through

the chip. Not only that, but the W-SWI-C also beats the VCT in area overhead at around 5 times less. Therefore, the W-SWI-D succeeds these architectures in terms of low area overheads and circuit complexity. However, W-SWI-C have slightly less area overhead than W-SWI-D.

#### 6.4 SUMMARY AND CONCLUSION

In this chapter we explored another design option to utilize 1-to-M communication, which is the decentralized approach to handle contentions. The stretch-multicast offers better flexibility on allocating and utilizing the SWI. Moreover, to enable the stretch-multicast while avoiding creating a deadlock situation, an efficient tag-based flow control has been proposed. Due to the merits of this approach in terms of SWI utilization, evaluation results show that W-SWI-D is better than W-SWI-C in terms of average delay and power consumption when the NoC architecture is virtual-channelless. Moreover, the comparison of the W-SWI-C and the W-SWI-D under multicast traffic has proven that the former is preferred for higher traffic loads while the latter is optimal for low traffic loads. In contrast, W-SWI-C has proven that it is suitable for high load or in case of broadcast traffic. These results, in general, demonstrate further improvements could be achieved on W-SWI based on the application requirement predictions and the design requirements. Therefore, the proposed hybrid W-SWI interconnects continue to show high potential and scalability for future NoC-based CMPs.

## CONCLUSIONS AND FUTURE WORK

---

### 7.1 SUMMARY AND CONCLUSION

The crave to implement more PE or IP in a single chip has expanded on-chip communication requirements to the degree that it has become a bottleneck for system performance. Moreover, this has created a paradigm shift in very large scale integration (VLSI) design from computational-centric to communicational-centric designs. This has been further amplified by continuous deterioration of the wire performance. Consequently, emerging technologies for interconnects are suggested to alleviate the growing challenges of on-chip communication. This thesis proposes the use of the SWI as the basis of a promising interconnect architecture, which can resolve two of the main existing challenges. These challenges are global and multicast communication. This section presents the main conclusions drawn from this thesis.

The problems with metal wires are projected to scale up as the technology scales down. The Zenneck surface wave seems to be a good alternative for intra-chip communication since it demonstrates promising characteristics such as low power dissipation, CMOS compatibility and high signal propagation speed. Moreover, to fully understand and analyse SWI, implementation considerations for SWI have been discussed, such as the needed integrated devices and RF communication channel design. In addition, the SWI system-level power modelling and performance considerations have been discussed based on real SWI experiments. The SWI experiments also confirm the superiority of SWI over wireless interconnects. This analysis and modelling have enabled a system-level evaluation, which is essential for highlighting potential improvements and challenges for the proposed architecture and motivate further investigation to reap the benefits of this interconnect technology.

A hybrid wire-SWI architecture is proposed, which efficiently exploits the SWI. The efficiency of this interconnect architecture is mainly due to the advantages of both the under-layer physical interconnects of SWI and the NoC's multilayer topology. Moreover, to efficiently utilize the proposed W-SWI architecture for global traffic, simple DWA algorithms are developed. Moreover, an implementation design along with an analysis of the area/power overhead of the extra circuitry required have been presented. Then, in order to explore the scalability potentials of the W-SWI in terms of global communication, a system-level evaluation has been conducted. The results show remarkable improvements in latency, throughput and power consumption without



any topology optimization algorithm or task mapping. All of this is achieved with only a relatively small die area penalty compared to baseline architecture. As a result, the proposed architecture enabled by SWI show promising potential for future on-chip unicast global communication.

Moreover, due to the necessity of multicast capabilities in many-core systems and the limitations of wire-based NoCs in this context, the hybrid W-SWI architecture has been considered for on-chip multicast communication. The potential of the proposed architecture to satisfy this demand has been enabled by the fact that the SWI shows extraordinary fan-out features compared to other emerging types of interconnect. These features enable the W-SWI to tackle 1-to-M traffic efficiently and gives an attractive area/performance trade-off. In addition, in order to resolve multicast contention issues, centralized (W-SWI-C) arbitration and allocation techniques are proposed and discussed. This includes designing a GMA, which shows high concurrent contention handling. Moreover, a tree-based multicast routing scheme for this architecture is suggested. The system-level evaluation shows substantial improvements in terms of average delay, saturated PIR and power consumption with a relatively small area overhead penalty compared to state-of-the-art on-chip multicast architectures. Consequently, the results demonstrate the high scalability of the W-SWI-C for future NoC-based many-core systems.

Furthermore, to explore the multicast contention handling, a decentralized arbitration and allocation scheme is proposed. This includes the development of stretch-multicast, which offers better flexibility in allocating and utilizing the SWI's shared channels by enabling slaves to be virtually allocated to more than one master. However, this risks the creation of multicast cycle dependency. Thus, to enable the stretch-multicast while avoiding creating deadlock situations, an efficient tag-based flow control method has been proposed. Due to the merits of this approach in terms of SWI utilization, the evaluation results show that the W-SWI-D is better than W-SWI-C in terms of average delay and power consumption when the NoC architecture design is virtual-channelless. However, the W-SWI-C is slightly better than W-SWI-D for higher traffic loads while the latter is ideal for low traffic loads. This is due to the increase in arbiter delay as the traffic load increases, which overcomes the SWI utilization gain achieved by W-SWI-D. In addition, it has been concluded that broadcast traffic can not utilize the the allocation flexibility offered by the W-SWI-D and thus decay the performance due to the drawbacks of per flit channel establishment latency. Consequently, the centralized approach performs better than the decentralized approach in the case of broadcast. However, both architectures are still superior compared to state-of-the-art multicast interconnect architectures.

## 7.2 FUTURE WORK

The objectives of this thesis include opening new research horizons in future intra-chip communication. Therefore, many research directions can be drawn and motivated from this thesis in terms of utilizing the SWI. These research directions can be classified based on the abstraction level they targeted in order to enable the realization and exploitation of this new emerging technology at application level, network and system level, electromagnetic and RF level, and communication level.

The application abstraction level includes mapping application tasks based on their traffic requirements and node accessibility to the SWI. Moreover, task migration might benefit from the new architecture's capabilities. In addition, cache coherence protocol messages might be adjusted based on the merits of the new architecture.

On the other hand, although this thesis has targeted the networking and system level, further research might be required in this abstract level. For instance, an investigation of the combination of 3D-NoC and SWI. This way the architecture might harvest the benefits of both of these emerging technologies. The contention handling scheme might be further extended to be globally-contention-aware by making the routing decision based on some cost function such as a contention status of all masters. On the other hand, the proposed architecture might evolve to be a dynamic topology where every node can switch between master/slave capability based on traffic monitoring. Even though these suggestions might increase the area overhead, the possible improvements resulted from factors, such as the system adaptability, could justify the extra cost.

In terms of the electromagnetic wave and RF-engineering level, further exploration and tuning of the design and materials of the surface and transducer at the die-level is required. This would then enable the integration, fabrication, and testing of the integrated on-chip surface and transducer. In addition, for the communication level, despite the fact that many integrated transceivers have been proposed in the literature with excellent characteristics, a transceiver that is specifically designed for the SWI might be more performance and/or energy efficient. On the other hand, optimizing the transceiver design and SNR using dynamic power scaling based on transmission distance, such as the transceiver proposed by Mineo et al. [93], would make the SWI further more efficient in terms of power consumption.

# Part II

## Thesis Appendices

## NOXIM SIMULATOR IMPROVEMENTS

---

This chapter presents the major added capabilities to the original Noxim simulator [121], in order to evaluate the proposed architecture and the developed techniques in this thesis. Although, Table 2.5 presents more added features, in here we will briefly discuss the simulator features that have not been discussed sufficiently in the main body of the thesis.

### A.1 SWI CHANNEL MODELLING

Modelling SWI channel in cycle-accurate simulator is simple enough since it is similar to modelling a memory block with one static write access right and multiple dynamic read access rights. These channels, as mentioned throughout the thesis, resemble a bus topology. The difficult part of system-level behaviour simulation is the modelling of the allocation/arbitration techniques and the communication protocols, which has been thoroughly discussed in this thesis. The adjustments to the Noxim simulator include defining SWI channels as global channels then give write privilege (Tx privilege) based on the input master placement matrix. The slave-read, however, is slightly more complicated and depends on the whether it is unicast or multicast and if multicast is it centralized or decentralized approach, see chapters 4, 5, and 6.

In the router module an extra port has to be added to the existing router ports: PE, E, W, N, and S, see Fig. A.1a. Therefore, also all the resulted implications of this extra port in concurrent process related to receiving and transmitting need to be adjusted. Some of the routing functions and selections functions such as odd-even has been adjusted to include the extra rules to route through and from the SWI. This extra port then linked to the SWI channel.

### A.2 VIRTUAL CHANNEL MODELLING

Noxim simulator does not support VC. Therefore, for the purpose of this study the Noxim has to be modified to simulate VC. In order to simulate the VC two components need to be altered, like in digital router design: The data-path and the control-plane. The data-path is altered by representing the buffer as a 2D storage (port  $\times$  VC) instead of one-dimension in the router module in Noxim, as shown in Fig. A.1a. This also would include nested iterations in the Tx process and

the Rx process to check and manage the First-In-First-Out (FIFO) of each VC in each port.

The control-plane alterations include the adjustment of the reservation table module in Noxim. This is also include storage space expansion from (input : port  $\times$  output : port) to (input : port, VC  $\times$  input : port, VC). Therefore, the reservation table module has to accept the input port and VC numbers and return the output port and VC numbers. Moreover, control signals need to be added in Tile and Router modules and then linked in the NoC module. These signals include a pair of Req/Ack for each VC in every port, as shown in Fig. A.1b.

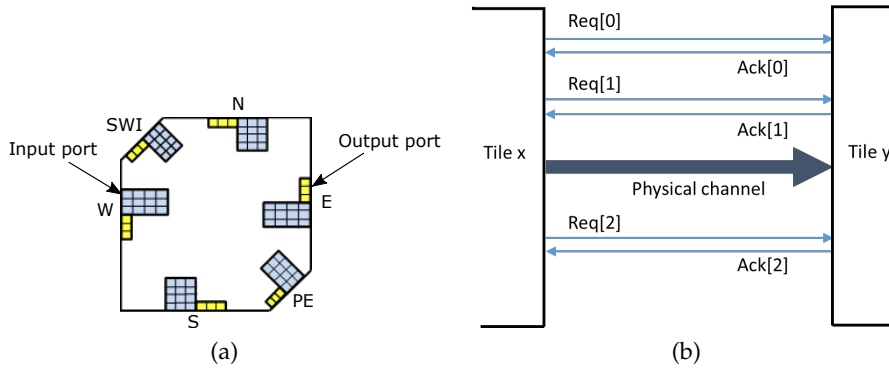


Figure A.1: (a) Demonstration of router ports with three VCs; (b) The added control signals between tiles to simulate three VCs in this example.

### A.3 1-TO-M TRAFFIC MODELLING

Although the Noxim simulator supports many synthetic traffic, they are all 1-to-1 traffic. Therefore, in order to illustrate the potentials of the proposed interconnect architecture in case of 1-to-M, the processing element module in Noxim has to be adjusted to include injection of multicast and broadcast traffics. These also need to be handled in either software-multicast, VCT, or W-SWI. The adopted technique to handle multicast traffic determine the way of injecting it.

Algorithm A.1 illustrate the adjustments to the processing element module in the Noxim simulator. The first part show how multicast or broadcast have been generated based on the input ratio. Then the destinations are selected randomly in case of multicast. The second part shows how to inject this multicast traffic into the NoC based on different interconnect architectures. Regarding the VCT details of injecting the traffic and then building the embedded multicast routing tree, the interested reader can found it in the literature [15].

---

**Algorithm A.1** Algorithm of the generation of 1-to-M or 1-to-all traffics and the way of injected them based on the architecture that handel the multicast-traffic; Where N is the number of nodes and Flit.M is the multicast flag bit.

---

.....**Multicast flit generation** -

**Input:** Multicast-Ratio, Broadcast/Multicast

```

1: X = Random( $0 \rightarrow 100$ )
2: if Multicast – Ratio  $\geq$  X then
3:   if Multicast then
4:     distillations = Random( $0 \rightarrow 2^N - 1$ )
5:     MAB = Binary(distillations)
6:     MAB[source] = 0
7:   else if Broadcastt then
8:     MAB = Binary( $2^N - 1$ )
9:     MAB[source] = 0
10:  end if
11: else
12:   Normal unicast...
13: end if

```

.....**Traffic injection** -

**Input:** Traffic handling Scheme: Software-Multicast/W-SWI/VCT

```

14: if Software – Multicast then
15:   for all  $i \in \{0, \dots, N - 1\}$  do
16:     if MAB[i] == 1 then
17:       Flit.destination = i
18:       Flit.M = 0
19:       Inject(Flit)
20:     end if
21:   end for
22: else if W – SWI then
23:   Flit.MAB = MAB
24:   Flit.M = 1
25:   Inject(Flit)
26: else if VCT then
27:   Either inject traffic in setup phase or multicast phase
28: end if

```

---

## Part III

# Thesis Bibliography

## BIBLIOGRAPHY

---

- [1] William J. Dally and John W. Poulton. *Digital systems engineering*. Cambridge University Press, New York, NY, USA, 1998.
- [2] W.J. Dally and B. Towles. Route packets, not wires: on-chip interconnection networks. In *Design Automation Conference, 2001. Proceedings*, pages 684 – 689, 2001.
- [3] J. Duato, S. Yalmanchili, and L. Ni. *Interconnection Networks : an Engineering Approach*. IEEE CS Press, Los Alamitos, Calif., 1997.
- [4] Tobias Bjerregaard and Shankar Mahadevan. A survey of research and practices of network-on-chip. *ACM Comput. Surv.*, 38(1), June 2006.
- [5] P. Salihundam, S. Jain, T. Jacob, S. Kumar, V. Erraguntla, Y. Hoskote, S. Vangal, G. Ruhl, and N. Borkar. A 2 tb/s 6  $\times$ , 4 mesh network for a single-chip cloud computer with dvfs in 45 nm cmos. *Solid-State Circuits, IEEE Journal of*, 46(4):757–766, April 2011.
- [6] S.R. Vangal, J. Howard, G. Ruhl, S. Dighe, H. Wilson, J. Tschanz, D. Finan, A. Singh, T. Jacob, S. Jain, V. Erraguntla, C. Roberts, Y. Hoskote, N. Borkar, and S. Borkar. An 80-tile sub-100-w teraflops processor in 65-nm cmos. *Solid-State Circuits, IEEE Journal of*, 43(1):29–41, Jan 2008.
- [7] Semiconductor Industry Association. ITRS: International Technology Roadmap for Semiconductors . <http://www.itrs.net/reports.html> [online], 2009.
- [8] R. Ho, K.W. Mai, and M.A. Horowitz. The future of wires. *Proceedings of the IEEE*, 89(4):490 –504, apr 2001.
- [9] Luca P. Carloni, Partha Pande, and Yuan Xie. Networks-on-chip in emerging interconnect paradigms: Advantages and challenges. In *Proceedings of the 2009 3rd ACM/IEEE International Symposium on Networks-on-Chip, NOCS '09*, pages 93–102, Washington, DC, USA, 2009. IEEE Computer Society.
- [10] Semiconductor Industry Association. Itrs: International technology roadmap for semiconductors. <http://www.itrs.net/reports.html> [online], 2012.
- [11] Shekhar Borkar. Thousand core chips: a technology perspective. In *Proceedings of the 44th annual Design Automation Conference, DAC '07*, pages 746–749, New York, NY, USA, 2007. ACM.



- [12] Daniel R. Johnson, Matthew R. Johnson, John H. Kelm, William Tuohy, Steven S. Lumetta, and Sanjay J. Patel. Rigel: A 1,024-core single-chip accelerator architecture. *IEEE Micro*, 31(4):30–41, 2011.
- [13] Simon Moore and Daniel Greenfield. The next resource war: computation vs. communication. In *Proceedings of the 2008 international workshop on System level interconnect prediction, SLIP '08*, pages 81–86, New York, NY, USA, 2008. ACM.
- [14] Semiconductor Industry Association. ITRS: International Technology Roadmap for Semiconductors . <http://www.itrs.net/reports.html> [online], 2005.
- [15] N.E. Jerger, Li-Shiuan Peh, and M. Lipasti. Virtual circuit tree multicasting: A case for on-chip hardware multicast support. In *Computer Architecture, 2008. ISCA '08. 35th International Symposium on*, pages 229–240, 2008.
- [16] Tushar Krishna, Li-Shiuan Peh, Bradford M. Beckmann, and Steven K. Reinhardt. Towards the ideal on-chip fabric for 1-to-many and many-to-1 communication. In *Proceedings of the 44th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO-44*, pages 71–82, NY, USA, 2011. ACM.
- [17] Sergi Abadal, Albert Mestres, Raul Martinez, Eduard Alarcon, Albert Cabellos-Aparicio, and Raul Martinez. Multicast on-chip traffic analysis targeting manycore noc design. In *Parallel, Distributed and Network-Based Processing (PDP), 2015 23rd Euromicro International Conference on*, pages 370–378, March 2015.
- [18] M. C F Chang, J. Cong, A. Kaplan, Chunyue Liu, M. Naik, J. Premkumar, G. Reinman, E. Socher, and Sai-Wang Tam. Power reduction of cmp communication networks via rf-interconnects. In *Microarchitecture, 2008. MICRO-41. 2008 41st IEEE/ACM International Symposium on*, pages 376–387, Nov 2008.
- [19] A. Carpenter, Jianyun Hu, Jie Xu, M. Huang, Hui Wu, and Peng Liu. Using transmission lines for global on-chip communication. *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, 2(2):183–193, June 2012.
- [20] B.M. Beckmann and D.A. Wood. Tlc: Transmission line caches. In *Microarchitecture, 2003. MICRO-36. Proceedings. 36th Annual IEEE/ACM International Symposium on*, pages 43–54, 2003.
- [21] Heng-Ming Hsu, Tai-Hsin Lee, and Chan-Jung Hsu. Millimeter-wave transmission line in 90-nm cmos technology. *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, 2(2):194–199, June 2012.

- [22] S. Deb, A Ganguly, P.P. Pande, B. Belzer, and D. Heo. Wireless noc as interconnection backbone for multicore chips: Promises and challenges. *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, 2(2):228–239, June 2012.
- [23] S. Abadal, M. Iannazzo, M. Nemirovsky, A. Cabellos-Aparicio, H. Lee, and E. Alarcon. On the area and energy scalability of wireless network-on-chip: A model-based benchmarked design space exploration. *Networking, IEEE/ACM Transactions on*, PP(99):1–1, 2014.
- [24] A. Ganguly, K. Chang, S. Deb, P.P. Pande, B. Belzer, and C. Teuscher. Scalable hybrid wireless network-on-chip architectures for multicore systems. *Computers, IEEE Transactions on*, 60(10):1485–1502, Oct 2011.
- [25] D. DiTomaso, A Kodi, S. Kaya, and D. Matolak. iwise: Inter-router wireless scalable express channels for network-on-chips (nocs) architecture. In *High Performance Interconnects (HOTI), 2011 IEEE 19th Annual Symposium on*, pages 11–18, Aug 2011.
- [26] Dan Zhao and Ruizhe Wu. Overlaid mesh topology design and deadlock free routing in wireless network-on-chip. In *Networks on Chip (NoCS), 2012 Sixth IEEE/ACM International Symposium on*, pages 27–34, May 2012.
- [27] N. Kirman, M. Kirman, R.K. Dokania, J.F. Martinez, A.B. Apsel, M.A. Watkins, and D.H. Albonesi. Leveraging optical technology in future bus-based chip multiprocessors. In *Microarchitecture, 2006. MICRO-39. 39th Annual IEEE/ACM International Symposium on*, pages 492–503, Dec 2006.
- [28] D. Miller. Device requirements for optical interconnects to silicon chips. *Proceedings of the IEEE*, 97(7):1166–1185, july 2009.
- [29] Po Dong, Young-Kai Chen, Tingyi Gu, Lawrence L. Buhl, David T. Neilson, and Jeffrey H. Sinsky. Reconfigurable 100 gb/s silicon photonic network-on-chip [invited]. *Optical Communications and Networking, IEEE/OSA Journal of*, 7(1):A37–A43, Jan 2015.
- [30] Dawei Huang, T. Sze, A. Landin, R. Lytel, and H.L. Davidson. Optical interconnects: out of the box forever? *Selected Topics in Quantum Electronics, IEEE Journal of*, 9(2):614–623, March 2003.
- [31] Yuan Liang, Hao Yu, Junfeng Zhao, Wei Yang, and Yuangang Wang. An energy efficient and low cross-talk cmos sub-thz i/o with surface-wave modulator and interconnect. In *Low Power Electronics and Design (ISLPED), 2015 IEEE/ACM International Symposium on*, pages 110–115, July 2015.

- [32] Michael Opoku Agyeman, Kenneth Tong, and Terrence Mak. An improved wireless communication fabric for emerging network-on-chip design. *Procedia Computer Science*, 56:415 – 420, 2015. The 10th International Conference on Future Networks and Communications (FNC 2015) / The 12th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2015) Affiliated Workshops.
- [33] Sergi Abadal, Benny Sheinman, Oded Katz, Ofer Markish, Danny Elad, Yvan Fournier, Damian Roca, Mauricio Hanzich, Guillaume Houzeaux, Mario Nemirovsky, Eduard Alarcon, and Albert Cabellos-Aparicio. Broadcast-enabled massive multicore architectures: A wireless rf approach. *Micro, IEEE*, 35(5):52–61, Sept 2015.
- [34] Ammar Karkar, Ra’ed Al-Dujaily, Alex Yakovlev, Kenneth Tong, and Terrence Mak. Surface wave communication system for on-chip and off-chip interconnects. In *Proceedings of the Fifth International Workshop on Network on Chip Architectures*, NoCArc ’12, pages 11–16, New York, NY, USA, 2012. ACM.
- [35] A.J. Karkar, J.E. Turner, K. Tong, R. Al-Dujaily, T. Mak, A. Yakovlev, and Fei Xia. Hybrid wire-surface wave interconnects for next-generation networks-on-chip. *Computers Digital Techniques, IET*, 7(6):294–303, November 2013.
- [36] A. Karkar, N. Dahir, R. Al-Dujaily, K. Tong, T. Mak, and A. Yakovlev. Hybrid wire-surface wave architecture for one-to-many communication in networks-on-chip. In *Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014*, pages 1–4, March 2014.
- [37] Ammar Karkar, Kin-Fai Tong, Terrence Mak, and Alex Yakovlev. Mixed wire and surface-wave communication fabrics for decentralized on-chip multicasting. In *Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition, DATE ’15*, pages 794–799, San Jose, CA, USA, 2015. EDA Consortium.
- [38] A. Karkar, T. Mak, K. Tong, and A. Yakovlev. A survey of emerging interconnects for on-chip efficient multicast and broadcast in many-cores. *Circuits and Systems Magazine, IEEE*, 16(1):58–72, Firstquarter 2016.
- [39] Ammar Karkar, Terrence Mak, and Alex Yakovlev. Demonstration of emerging on-chip surface-wave interconnects. In *(Submitted), 10th IEEE/ACM International Symposium on Networks-on-Chip, NOCS ’16*, 2016.
- [40] A. Karkar, T. Mak, N. Dahir, R. Al-Dujaily, K. F. Tong, and A. Yakovlev. Network-on-chip multicast architectures using

- hybrid wire and surface-wave interconnects. *IEEE Transactions on Emerging Topics in Computing*, PP(99):1–1, 2016.
- [41] Partha Pratim Pande, C. Grecu, M. Jones, A. Ivanov, and R. Saleh. Performance evaluation and design trade-offs for network-on-chip interconnect architectures. *Computers, IEEE Transactions on*, 54(8):1025 – 1040, aug. 2005.
- [42] R. Marculescu, Jingcao Hu, and U.Y. Ogras. Key research problems in NoC design: a holistic perspective. In *Hardware/Software Codesign and System Synthesis, 2005. CODES+ISSS '05. Third IEEE/ACM/IFIP International Conference on*, pages 69 –74, sept. 2005.
- [43] R. Marculescu, U.Y. Ogras, Li-Shiuan Peh, N.E. Jerger, and Y. Hoskote. Outstanding research problems in noc design: System, microarchitecture, and circuit perspectives. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 28(1):3 –21, jan. 2009.
- [44] L. Benini and G. De Micheli. Networks on chips: a new soc paradigm. *IEEE Computer*, 35(1):70 –78, 2002.
- [45] S. Dighe, S.R. Vangal, P. Aseron, S. Kumar, T. Jacob, K.A. Bowman, J. Howard, J. Tschanz, V. Erraguntla, N. Borkar, V.K. De, and S. Borkar. Within-die variation-aware dynamic-voltage-frequency-scaling with optimal core allocation and thread hopping for the 80-core teraflops processor. *Solid-State Circuits, IEEE Journal of*, 46(1):184 –193, jan. 2011.
- [46] Tiler. Tilepro processor family, Dec 2011.
- [47] M.B. Taylor, J. Kim, J. Miller, D. Wentzlaff, F. Ghodrat, B. Greenwald, H. Hoffman, P. Johnson, Jae-Wook Lee, W. Lee, A. Ma, A. Saraf, M. Seneski, N. Shnidman, V. Strumpfen, M. Frank, S. Amarasinghe, and A. Agarwal. The raw microprocessor: a computational fabric for software circuits and general-purpose programs. *Micro, IEEE*, 22(2):25–35, 2002.
- [48] Terrence Sui-Tung Mak. *Circuit design and analysis for on-FPGA communication systems*. PhD thesis, Imperial College London, 2009.
- [49] ARM. Amba specification (rev 2.0). 1999.
- [50] Sungju Han, Jinho Lee, and Kiyoun Choi. Tree-mesh heterogeneous topology for low-latency noc. In *Proceedings of the 2014 International Workshop on Network on Chip Architectures, NoCArc '14*, pages 19–24, New York, NY, USA, 2014. ACM.

- [51] William Dally and Brian Towles. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann, 2004.
- [52] Ra'ed Al-Dujaily. Embedded dynamic programming networks for networks-on-chip. 2013.
- [53] W.J. Dally. Virtual-channel flow control. *Parallel and Distributed Systems, IEEE Transactions on*, 3(2):194–205, Mar 1992.
- [54] Christopher J. Glass and Lionel M. Ni. The turn model for adaptive routing. *SIGARCH Comput. Archit. News*, 20(2):278–287, April 1992.
- [55] Ge-Ming Chiu. The odd-even turn model for adaptive routing. *Parallel and Distributed Systems, IEEE Transactions on*, 11(7):729–738, jul 2000.
- [56] N. Dahir, T. Mak, R. Al-Dujaily, and Yakovlev A. Highly adaptive and deadlock-free routing for three-dimensional networks-on-chip. *IET Computers and Digital Techniques*, 7:255–263(8), November 2013.
- [57] K. Jain, S.K. Singh, A. Majumder, and A.J. Mondai. Problems encountered in various arbitration techniques used in noc router: A survey. In *Electronic Design, Computer Networks Automated Verification (EDCAV), 2015 International Conference on*, pages 62–67, Jan 2015.
- [58] Shekhar Borkar and Andrew A. Chien. The future of microprocessors. *Commun. ACM*, 54(5):67–77, May 2011.
- [59] G. Blake, R.G. Dreslinski, and T. Mudge. A survey of multi-core processors. *Signal Processing Magazine, IEEE*, 26(6):26–37, November 2009.
- [60] Christian Martin. Multicore processors: Challenges, opportunities, emerging trends. In *Proceedings of the 2014 Embedded world conference*, 2014.
- [61] Neil HE Weste, David Harris, et al. *CMOS VLSI design: a circuits and systems perspective*. Boston: Pearson/Addison-Wesley, 2005.
- [62] Daniel J Sorin, Mark D Hill, and David A Wood. A primer on memory consistency and cache coherence. *Synthesis Lectures on Computer Architecture*, 6(3):1–212, 2011.
- [63] Limin Han, Jianfeng An, Deyuan Gao, Xiaoya Fan, Xianglong Ren, and Tao Yao. A survey on cache coherence for tiled many-core processor. In *Signal Processing, Communication and Computing (ICSPCC), 2012 IEEE International Conference on*, pages 114–118, Aug 2012.

- [64] V Pavlidis and Eby G Friedman. *Three-dimensional integrated circuit design*. Morgan Kaufmann, 2010.
- [65] Hiroki Matsutani, Michihiro Koibuchi, Tadahiro Kuroda, and Hideharu Amano. 3-d noc on inductive wireless interconnect. pages 225–248, 2011.
- [66] V.F. Pavlidis and E.G. Friedman. 3-d topologies for networks-on-chip. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 15(10):1081–1090, Oct 2007.
- [67] Nizar Dahir. Physical parameter-aware networks-on-chip design. 2015.
- [68] Marion Rey F. Abingosa, Clemente Receno, John Imperial, and Jefferson A. Hora. Interconnect modeling of global metals for 40nm node. In *Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 2015 International Conference on*, pages 1–6, Dec 2015.
- [69] K. Banerjee and A. Mehrotra. A power-optimal repeater insertion methodology for global interconnects in nanometer designs. *Electron Devices, IEEE Transactions on*, 49(11):2001 – 2007, nov 2002.
- [70] K. Ohashi, K. Nishi, T. Shimizu, M. Nakada, J. Fujikata, J. Ushida, S. Torii, K. Nose, M. Mizuno, H. Yukawa, M. Kinoshita, N. Suzuki, A. Gomyo, T. Ishi, D. Okamoto, K. Furue, T. Ueno, T. Tsuchizawa, T. Watanabe, K. Yamada, S. Itabashi, and J. Akedo. On-chip optical interconnect. *Proceedings of the IEEE*, 97(7):1186–1198, July 2009.
- [71] R. Meade, J.S. Orcutt, K. Mehta, O. Tehar-Zahav, D. Miller, M. Georgas, B. Moss, Chen Sun, Yu-Hsin Chen, J. Shainline, M. Wade, R. Bafrali, Z. Sternberg, G. Machavariani, G. Sandhu, M. Popovic, R. Ram, and V. Stojanovic. Integration of silicon photonics in bulk cmos. In *VLSI Technology (VLSI-Technology): Digest of Technical Papers, 2014 Symposium on*, pages 1–2, June 2014.
- [72] J.E. Cunningham, I. Shubin, H.D. Thacker, Jin-Hyoung Lee, Guoliang Li, Xuezhe Zheng, J. Lexau, R. Ho, J.G. Mitchell, Ying Luo, Jin Yao, K. Raj, and A.V. Krishnamoorthy. Scaling hybrid-integration of silicon photonics in freescale 130nm to tsmc 40nm-cmos vlsi drivers for low power communications. In *Electronic Components and Technology Conference (ECTC), 2012 IEEE 62nd*, pages 1518–1525, May 2012.
- [73] Xuezhe Zheng, Dinesh Patil, Jon Lexau, Frankie Liu, Guoliang Li, Hiren Thacker, Ying Luo, Ivan Shubin, Jieda Li, Jin Yao, Po Dong,

- Dazeng Feng, Mehdi Asghari, Thierry Pinguet, Attila Mekis, Philip Amberg, Michael Dayringer, Jon Gainsley, Hesam Fathi Moghadam, Elad Alon, Kannan Raj, Ron Ho, John E. Cunningham, and Ashok V. Krishnamoorthy. Ultra-efficient 10gb/s hybrid integrated silicon photonic transmitter and receiver. *Opt. Express*, 19(6):5172–5186, Mar 2011.
- [74] S. Faralli, F. Gambini, P. Pintus, O. Liboiron-Ladouceur, P. Castoldi, N. Andriolli, and I. Cerutti. Experimental demonstration of bidirectional transmissions in a photonic integrated network on chip with bus topology. In *Photonics in Switching (PS), 2015 International Conference on*, pages 357–359, Sept 2015.
- [75] Jing Xue, Alok Garg, Berkehan Ciftcioglu, Jianyun Hu, Shang Wang, Ioannis Savidis, Manish Jain, Rebecca Berman, Peng Liu, Michael Huang, Hui Wu, Eby Friedman, Gary Wicks, and Duncan Moore. An intra-chip free-space optical interconnect. *SIGARCH Comput. Archit. News*, 38(3):94–105, June 2010.
- [76] Rajeev K. Dokania and Alyssa B. Apsel. Analysis of challenges for on-chip optical interconnects. In *Proceedings of the 19th ACM Great Lakes Symposium on VLSI, GLSVLSI '09*, pages 275–280, New York, NY, USA, 2009. ACM.
- [77] M. Haurylau, Hui Chen, Jidong Zhang, Guoqing Chen, N.A. Nelson, D.H. Albonesi, E.G. Friedman, and P.M. Fauchet. On-chip optical interconnect roadmap: challenges and critical directions. pages 17–19, Sept 2005.
- [78] A.R. Mickelson. Silicon photonics for on-chip interconnections. In *Custom Integrated Circuits Conference (CICC), 2011 IEEE*, pages 1–8, Sept 2011.
- [79] M. Mohamed, Zheng Li, Xi Chen, Li Shang, and A.R. Mickelson. Reliability-aware design flow for silicon photonics on-chip interconnect. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 22(8):1763–1776, Aug 2014.
- [80] R.G. Kim, W. Choi, Z. Chen, P.P. Pande, D. Marculescu, and R. Marculescu. Wireless noc and dynamic vfi codesign: Energy efficiency without performance penalty. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, PP(99):1–14, 2016.
- [81] K.K. O, K. Kim, B. Floyd, J. Mehta, H. Yoon, C.-M. Hung, D. Bravo, T. Dickson, X. Guo, R. Li, N. Trichy, J. Caserta, W. Bomstad, J. Branch, D.-J. Yang, J. Bohorquez, J. Chen, E.-Y. Seok, L. Gao, A. Sugavanam, J.-J. Lin, S. Yu, C. Cao, M.-H. Hwang, Y.-R. Ding, S.-H. Hwang, H. Wu, N. Zhang, and J.E. Brewer. The feasibility of on-chip interconnection using antennas. In

*Computer-Aided Design, 2005. ICCAD-2005. IEEE/ACM International Conference on*, pages 979–984, Nov 2005.

- [82] M.F. Chang, J. Cong, A. Kaplan, M. Naik, G. Reinman, E. Socher, and S.-W. Tam. Cmp network-on-chip overlaid with multi-band rf-interconnect. In *High Performance Computer Architecture, 2008. HPCA 2008. IEEE 14th International Symposium on*, pages 191–202, Feb 2008.
- [83] Daquan Huang, T.R. LaRocca, L. Samoska, A. Fung, and M.-C.F. Chang. 324ghz cmos frequency generator using linear superposition technique. In *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, pages 476–629, Feb 2008.
- [84] Wu-Hsin Chen, Sanghoon Joo, S. Sayilir, R. Willmot, Tae-Young Choi, Dowon Kim, J. Lu, D. Peroulis, and Byunghoo Jung. A 6-gb/s wireless inter-chip data link using 43-ghz transceivers and bond-wire antennas. *Solid-State Circuits, IEEE Journal of*, 44(10):2711–2721, Oct 2009.
- [85] Huaide Wang, Meng-Hsiung Hung, Yu-Ching Yeh, and Jri Lee. A 60-ghz fsk transceiver with automatically-calibrated demodulator in 90-nm cmos. In *VLSI Circuits (VLSIC), 2010 IEEE Symposium on*, pages 95–96, June 2010.
- [86] Xinmin Yu, S.P. Sah, S. Deb, P.P. Pande, B. Belzer, and Deukhyoun Heo. A wideband body-enabled millimeter-wave transceiver for wireless network-on-chip. In *Circuits and Systems (MWSCAS), 2011 IEEE 54th International Midwest Symposium on*, pages 1–4, Aug 2011.
- [87] K. Kawasaki, Y. Akiyama, K. Komori, M. Uno, H. Takeuchi, T. Itagaki, Y. Hino, Y. Kawasaki, K. Ito, and A. Hajimiri. A millimeter-wave intra-connect solution. *Solid-State Circuits, IEEE Journal of*, 45(12):2655–2666, Dec 2010.
- [88] K. Okada, K. Kondou, M. Miyahara, M. Shinagawa, H. Asada, R. Minami, T. Yamaguchi, A. Musa, Y. Tsukui, Y. Asakura, S. Tamonoki, H. Yamagishi, Y. Hino, T. Sato, H. Sakaguchi, N. Shimasaki, T. Ito, Y. Takeuchi, N. Li, Q. Bu, R. Murakami, K. Bunsen, K. Matsushita, M. Noda, and A. Matsuzawa. Full four-channel 6.3-gb/s 60-ghz cmos transceiver with low-power analog and digital baseband circuitry. *Solid-State Circuits, IEEE Journal of*, 48(1):46–65, Jan 2013.
- [89] S. Kawai, R. Minami, Y. Tsukui, Y. Takeuchi, H. Asada, A. Musa, R. Murakami, T. Sato, Qinghong Bu, Ning Li, M. Miyahara, K. Okada, and A. Matsuzawa. A digitally-calibrated 20-gb/s 60-ghz direct-conversion transceiver in 65-nm cmos. In *Radio*



- Frequency Integrated Circuits Symposium (RFIC), 2013 IEEE*, pages 137–140, June 2013.
- [90] Jau Lin, Hsin-Ta Wu, Yu Su, Li Gao, A. Sugavanam, J.E. Brewer, and K.K. O. Communication using antennas fabricated in silicon integrated circuits. *Solid-State Circuits, IEEE Journal of*, 42(8):1678–1687, Aug 2007.
  - [91] Xinmin Yu, S.P. Sah, H. Rashtian, S. Mirabbasi, P.P. Pande, and Deukhyoun Heo. A 1.2-pj/bit 16-gb/s 60-ghz ook transmitter in 65-nm cmos for wireless network-on-chip. *Microwave Theory and Techniques, IEEE Transactions on*, 62(10):2357–2369, Oct 2014.
  - [92] B.A Floyd, Chih-Ming Hung, and K.K. O. Intra-chip wireless interconnect for clock distribution implemented with integrated antennas, receivers, and transmitters. *Solid-State Circuits, IEEE Journal of*, 37(5):543–552, May 2002.
  - [93] A Mineo, M. Palesi, G. Ascia, and V. Catania. An adaptive transmitting power technique for energy efficient mm-wave wireless nocs. In *Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014*, pages 1–6, March 2014.
  - [94] A. Ganguly, K. Chang, S. Deb, P.P. Pande, and *et al.* Scalable hybrid wireless network-on-chip architectures for multicore systems. *Computers, IEEE Transactions on*, 60(10):1485–1502, Oct 2011.
  - [95] Yi Huang, Wen-Yan Yin, and Qing Huo Liu. Performance prediction of carbon nanotube bundle dipole antennas. *Nanotechnology, IEEE Transactions on*, 7(3):331–337, May 2008.
  - [96] J.L. Bohorquez and O. Kenneth. A study of the effects of microwave electromagnetic radiation on dynamic random access memory operation. In *Electromagnetic Compatibility. EMC 2004. International Symposium on*, volume 3, pages 815–819, Aug 2004.
  - [97] M.H. Hwang and O. Kenneth. A study of the impact of microwave radiation on an a/d converter. In *Electromagnetic Compatibility, 2005. EMC 2005. 2005 International Symposium on*, volume 2, pages 307–311 Vol. 2, Aug 2005.
  - [98] M.-C.F. Chang, V.P. Roychowdhury, Liyang Zhang, Hyunchol Shin, and Yongxi Qian. Rf/wireless interconnect for inter- and intra-chip communications. *Proceedings of the IEEE*, 89(4):456–466, Apr 2001.
  - [99] J.D. Warnock, J.M. Keaty, J. Petrovick, J.G. Clabes, C.J. Kircher, B.L. Krauter, P.J. Restle, B.A. Zoric, and C.J. Anderson. The circuit and physical design of the power4 microprocessor. *IBM Journal of Research and Development*, 46(1):27–51, Jan 2002.

- [100] H. Ito, M. Kimura, K. Miyashita, T. Ishii, K. Okada, and K. Masu. A bidirectional- and multi-drop-transmission-line interconnect for multipoint-to-multipoint on-chip communications. *Solid-State Circuits, IEEE Journal of*, 43(4):1020–1029, April 2008.
- [101] M-C Frank Chang, Eran Socher, Sai-Wang Tam, Jason Cong, and Glenn Reinman. Rf interconnects for communications on-chip. In *Proceedings of the 2008 international symposium on Physical design, ISPD '08*, pages 78–83, New York, NY, USA, 2008. ACM.
- [102] Terence Charles Edwards and Michael Bernard Steer. *Foundations of interconnect and microstrip design*, volume 3. John Wiley, 2000.
- [103] Suk-Bok Lee, Sai-Wang Tam, Ioannis Pefkianakis, Songwu Lu, M. Frank Chang, Chuanxiong Guo, Glenn Reinman, Chunyi Peng, Mishali Naik, Lixia Zhang, and Jason Cong. A scalable micro wireless interconnect structure for cmps. In *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking, MobiCom '09*, pages 217–228, New York, NY, USA, 2009. ACM.
- [104] Aaron Carpenter. *The design and use of high-speed transmission line links for global on-chip communication*. PhD thesis, University of Rochester, 2012.
- [105] J. Hendry. Isolation of the zenneck surface wave. In *Antennas and Propagation Conference (LAPC), 2010 Loughborough*, pages 613–616, nov. 2010.
- [106] J. Turner, M. Jessup, and K. Tong. A novel technique enabling the realisation of 60 GHz body area networks. In *Wearable and Implantable Body Sensor Networks (BSN), 2012 Ninth International Conference on*, pages 58–62, may 2012.
- [107] J. Hendry and M Underhill. Surface waves for communication systems. In *3rd SEAS DTC Technical Conference, Edinburgh 2008*. CC012, Ref, 2008.
- [108] JiXiang Wan, Kin Fai Tong, and ChunBang Wu. The excitation efficiency of surface waves on a reactive surface by a finite vertical aperture. In *Antennas and Propagation USNC/URSI National Radio Science Meeting, 2015 IEEE International Symposium on*, pages 1634–1635, July 2015.
- [109] David M Pozar. *Microwave Engineering*. John Wiley & Sons, 2009.
- [110] ANSYS. Hfss: high frequency simulation software. <http://www.ansys.com/Products/Electronics/ANSYS-HFSS> [online].

- [111] CST. CST: high computer simulation technology. <https://www.cst.com/> [online].
- [112] M.O. Agyeman, Kin-Fai Tong, and T. Mak. Towards reliability and performance-aware wireless network-on-chip design. In *Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFTS), 2015 IEEE International Symposium on*, pages 205–210, Oct 2015.
- [113] N. Magen, A. Kolodny, U. Weiser, and N. Shamir. Interconnect-power dissipation in a microprocessor. In *International Workshop on System Level Interconnect Prediction, SLIP*, pages 7–13, 2004. Cited By (since 1996): 128.
- [114] SpiNNaker home page. [Feb. 14, 2012].
- [115] S.B. Furber, F. Galluppi, S. Temple, and L.A. Plana. The spinnaker project. *Proceedings of the IEEE*, 102(5):652–665, May 2014.
- [116] S.B. Furber, D.R. Lester, L.A. Plana, J.D. Garside, E. Painkras, S. Temple, and A.D. Brown. Overview of the spinnaker system architecture. *Computers, IEEE Transactions on*, 62(12):2454–2467, 2013.
- [117] Javier Navaridas, Mikel Luján, Jose Miguel-Alonso, Luis A. Plana, and Steve Furber. Understanding the interconnection network of spinnaker. In *Proceedings of the 23rd International Conference on Supercomputing, ICS '09*, pages 286–295, New York, NY, USA, 2009. ACM.
- [118] James Jeffers and James Reinders. *Intel Xeon Phi coprocessor high-performance programming*. Intel Press, 2013.
- [119] Intel. Many Integrated Core (MIC) Architecture. [March. 12, 2012].
- [120] BookSim: Interconnection Network Simulator. [March 20, 2012].
- [121] F. Fazzino, M. Palesi, and D. Patti. Noxim: Network-on-chip simulator.
- [122] A.B. Kahng, Bin Li, Li-Shiuan Peh, and K. Samadi. Orion 2.0: A power-area simulator for interconnection networks. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 20(1):191–196, Jan 2012.
- [123] A.B. Kahng, Bill Lin, and S. Nath. Explicit modeling of control and data for improved noc router estimation. In *Design Automation Conference (DAC), 2012 49th ACM/EDAC/IEEE*, pages 392–397, June 2012.

- [124] R. Mullins. Netmaker: Interconnection Network Simulator. [April 26, 2012].
- [125] L. Jain, B. Al-Hashimi, M Zwolinski, M. Gaur, P. Rosinger, and V. Laxmi. MIRGAM: a simulator for NoC interconnect routing and application modeling. [April 26, 2012].
- [126] J. Chan, G. Hendry, A. Biberman, K. Bergman, and L.P. Carloni. Phoenixsim: A simulator for physical-layer analysis of chip-scale photonic interconnection networks. In *Design, Automation Test in Europe Conference Exhibition (DATE), 2010*, pages 691–696, March 2010.
- [127] A. Mineo, M. Palesi, G. Ascia, and V. Catania. Runtime tunable transmitting power technique in mm-wave winoc architectures. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, PP(99):1–1, 2015.
- [128] Vincenzo Catania, Andrea Mineo, Salvatore Monteleone, Maurizio Palesi, and Davide Patti. Noxim: An open, extensible and cycle-accurate network on chip simulator. In *Application-specific Systems, Architectures and Processors (ASAP), 2015 IEEE 26th International Conference on*, pages 162–163, July 2015.
- [129] K. Swaminathan, D. Thakyal, S.G. Nambiar, G. Lakshminarayanan, and Seok-Bum Ko. Enhanced noxim simulator for performance evaluation of network on chip topologies. In *Engineering and Computational Sciences (RAECS), 2014 Recent Advances in*, pages 1–5, March 2014.
- [130] L. Schares, J.A. Kash, F.E. Doany, C.L. Schow, C. Schuster, D.M. Kuchta, P.K. Pepeljugoski, J.M. Trehwella, C.W. Baks, R.A. John, Lei Shan, Y.H. Kwark, R.A. Budd, P. Chiniwalla, F.R. Libsch, J. Rosner, C.K. Tsang, G.S. Patel, J.D. Schaub, R. Dangel, F. Horst, B.J. Offrein, D. Kucharski, D. Guckenberger, S. Hegde, H. Nyikal, Chao-Kun Lin, Ashish Tandon, G.R. Trott, M. Nystrom, D.P. Bour, M.R.T. Tan, and D.W. Dolfi. Terabus: Terabit/second-class card-level optical interconnect technologies. *Selected Topics in Quantum Electronics, IEEE Journal of*, 12(5):1032–1044, Sept 2006.
- [131] Paul H Mazurkiewicz. Board-level conformal emi shield having an electrically-conductive polymer coating over a thermally-conductive dielectric coating, February 1 2005. US Patent 6,849,800.
- [132] Xin Lu and Gu Xu. Thermally conductive polymer composites for electronic packaging. *Journal of applied polymer science*, 65(13):2733–2738, 1997.

- [133] T.Y. Wu, S.W. Chua, and Y.L. Lu. Noise floor and dynamic range analysis of a microwave attenuation measurement receiver from 50 mhz to 26.5 ghz. *Measurement*, 44(9):1516 – 1525, 2011.
- [134] U.Y. Ogras and R. Marculescu. Application-specific network-on-chip architecture customization via long-range link insertion. In *Computer-Aided Design, 2005. ICCAD-2005. IEEE/ACM International Conference on*, pages 246–253, Nov 2005.
- [135] A Kumar, Li-Shiuan Peh, P. Kundu, and N.K. Jha. Toward ideal on-chip communication using express virtual channels. *Micro, IEEE*, 28(1):80–90, Jan 2008.
- [136] M.B. Stensgaard and J. Sparso. Renoc: A network-on-chip architecture with reconfigurable topology. In *Networks-on-Chip, 2008. NoCS 2008. Second ACM/IEEE International Symposium on*, pages 55–64, April 2008.
- [137] B. Grot, J. Hestness, S.W. Keckler, and O. Mutlu. Express cube topologies for on-chip interconnects. In *High Performance Computer Architecture, 2009. HPCA 2009. IEEE 15th International Symposium on*, pages 163–174, Feb 2009.
- [138] Hsin-Chou Chi and Chih-Tsung Tang. A deadlock-free routing scheme for interconnection networks with irregular topologies. In *Parallel and Distributed Systems, 1997. Proceedings., 1997 International Conference on*, pages 88–95, Dec 1997.
- [139] O. Lysne, T. Skeie, S.-A. Reinemo, and I. Theiss. Layered routing in irregular networks. *Parallel and Distributed Systems, IEEE Transactions on*, 17(1):51–65, Jan 2006.
- [140] P. Wettin, R. Kim, J. Murray, Xinmin Yu, P.P. Pande, A. Ganguly, and D. Heoamlan. Design space exploration for wireless nocs incorporating irregular network routing. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 33(11):1732–1745, Nov 2014.
- [141] K. Sekar, K. Lahiri, A. Raghunathan, and S. Dey. Dynamically configurable bus topologies for high-performance on-chip communication. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 16(10):1413–1426, Oct 2008.
- [142] U.Y. Ogras and R. Marculescu. It’s a small world after all: Noc performance optimization via long-range link insertion. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 14(7):693 –706, july 2006.
- [143] Mingmin Yuan, Weiwei Fu, Tianzhou Chen, and Minghui Wu. An exploration on quantity and layout of wireless nodes for

- hybrid wireless network-on-chip. In *High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC,CSS,ICSS), 2014 IEEE Intl Conf on*, pages 100–107, Aug 2014.
- [144] Mengyuan Wu, A. Karkar, Bo Liu, A. Yakovlev, G. Gielen, and V. Grout. Network on chip optimization based on surrogate model assisted evolutionary algorithms. In *Evolutionary Computation (CEC), 2014 IEEE Congress on*, pages 3266–3271, July 2014.
  - [145] Bo Liu, Francisco V Fernández, Georges Gielen, Ammar Karkar, Alex Yakovlev, and Vic Grout. Smas: A generalized and efficient framework for computationally expensive electronic design optimization problems. *Computational Intelligence in Analog and Mixed-Signal (AMS) and Radio-Frequency (RF) Circuit Design*, page 251, 2015.
  - [146] Nizar Dahir and *et al.* Minimizing power supply noise through harmonic mappings in networks-on-chip. In *Proceedings of the Eighth IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis, CODES+ISSS '12*, pages 113–122, New York, NY, USA, 2012. ACM.
  - [147] A.B. Kahng, Bin Li, Li-Shiuan Peh, and K. Samadi. Orion 2.0: A fast and accurate noc power and area model for early-stage design space exploration. In *Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09.*, pages 423–428, April 2009.
  - [148] J. Srinivasan, S.V. Adve, P. Bose, and J.A. Rivers. The impact of technology scaling on lifetime reliability. In *Dependable Systems and Networks, 2004 International Conference on*, pages 177 – 186, june-1 july 2004.
  - [149] P. Conway, N. Kalyanasundharam, G. Donley, K. Lepak, and B. Hughes. Cache hierarchy and memory subsystem of the amd opteron processor. *Micro, IEEE*, 30(2):16–29, March 2010.
  - [150] Christian Bienia, Sanjeev Kumar, Jaswinder Pal Singh, and Kai Li. The parsec benchmark suite: Characterization and architectural implications. In *Proceedings of the 17th International Conference on Parallel Architectures and Compilation Techniques, PACT '08*, pages 72–81, New York, NY, USA, 2008. ACM.
  - [151] Jaswinder Pal Singh, Wolf-Dietrich Weber, and Anoop Gupta. Splash: Stanford parallel applications for shared-memory. *SIGARCH Comput. Archit. News*, 20(1):5–44, March 1992.

- [152] Jian Wu and Steve Furber. A multicast routing scheme for a universal spiking neural network architecture. *Comput. J.*, 53(3):280–288, March 2010.
- [153] S. Carrillo, J. Harkin, L.J. McDaid, F. Morgan, S. Pande, S. Cawley, and B. McGinley. Scalable hierarchical network-on-chip architecture for spiking neural network hardware implementations. *Parallel and Distributed Systems, IEEE Transactions on*, 24(12):2451–2461, Dec 2013.
- [154] Xiaola Lin, P.K. McKinley, and L.M. Ni. Deadlock-free multicast wormhole routing in 2-d mesh multicomputers. *Parallel and Distributed Systems, IEEE Transactions on*, 5(8):793–804, Aug 1994.
- [155] F. Samman, T. Hollstein, and M. Glesner. New theory for deadlock-free multicast routing in wormhole-switched virtual-channelless networks-on-chip. *Parallel and Distributed Systems, IEEE Transactions on*, 22(4):544–557, April 2011.
- [156] Faizal Arya Samman and Thomas Hollstein. Efficient and deadlock-free tree-based multicast routing methods for networks-on-chip (noc). In Maurizio Palesi and Masoud Daneshatalab, editors, *Routing Algorithms in Networks-on-Chip*, pages 129–159. Springer New York, 2014.
- [157] P. Bahrebar and D. Stroobandt. Improving hamiltonian-based routing methods for on-chip networks: A turn model approach. In *Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014*, pages 1–4, March 2014.
- [158] M. Ebrahimi, M. Daneshatalab, P. Liljeberg, and H. Tenhunen. Hamum - a novel routing protocol for unicast and multicast traffic in mpsoes. In *Parallel, Distributed and Network-Based Processing (PDP), 2010 18th Euromicro International Conference on*, pages 525–532, Feb 2010.
- [159] R. Prolonge, F. Clermidy, L. Tedesco, and F. Moraes. Dynamic flow reconfiguration strategy to avoid communication hot-spots. In *Digital System Design (DSD), 2011 14th Euromicro Conference on*, pages 519–524, 2011.
- [160] R. Morris, E. Jolley, and A. Karanth Kodi. Extending the performance and energy-efficiency of shared memory multicores with nanophotonic technology, 2013.
- [161] D. DiTomaso, A. Kodi, D. Matolak, S. Kaya, S. Laha, and W. Rayess. A-winoc: Adaptive wireless network-on-chip architecture for chip multiprocessors. *Parallel and Distributed Systems, IEEE Transactions on*, PP(99):1–1, 2014.

- [162] T. Binzegger, R. Douglas, and K. Martin. A quantitative map of the circuit of cat primary visual cortex. *The Journal of Neuroscience*, 24:8441–8453.
- [163] George Karypis and Vipin Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM J. Sci. Comput.*, 20(1):359–392, December 1998.
- [164] D Kinniment. *Synchronization, Arbitration and Choice*. Wiley Publishing, England, 2007.
- [165] Tom Shanley. *Pentium Pro Processor System Architecture*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1996.